The Dissertation Committee for Marisa Elena Miller certifies that this is the approved version of the following dissertation:

Roles for Polyploidy, Circadian Rhythms, and Stress Responses in Hybrid Vigor

Committee:

Z. Jeffrey Chen, Supervisor

Sibum Sung

Enamul Huq

Alan Lloyd

Thomas Juenger

Roles for Polyploidy, Circadian Rhythms, and Stress Responses in Hybrid Vigor

by

Marisa Elena Miller, B.S. Bio.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2014

Acknowledgements

First, I would like to express my endless gratitude for my husband Logan and my family. Without their love and support, the desire to pursue my dream of becoming a scientist would be sorely lacking. In the words of Carl Sagan, “For small creatures such as we the vastness is bearable only through love.” I would also like to thank my advisor Dr. Z. Jeffrey Chen for his guidance. He taught me how to develop my scientific reasoning, and always provided me with many opportunities to improve my skills and expertise. My committee members, Dr. Sibum Sung, Dr. Alan Lloyd, Dr. Enamul Huq, and Dr. Thomas Juenger, also deserve a special mention for being wonderful role models, as well as for their excellent comments and critiques of my research. Lastly, I would like to thank all past and present members of the Chen lab for the many thoughtful discussions and good times we shared together. Their insights, advice, and friendship were invaluable.

Roles for Polyploidy, Circadian Rhythms, and Stress Responses in Hybrid Vigor

Marisa Elena Miller, PhD

The University of Texas at Austin, 2014

Supervisor: Z. Jeffrey Chen

Hybrid plants and animals, like corn and the domestic dog, grow larger and more vigorously than their parents, a common phenomenon known as hybrid vigor or heterosis. In hybrids between *Arabidopsis* ecotypes or species (in allotetraploids), altered expression of circadian clock genes leads to increased starch and chlorophyll content and greater biomass. In plants and animals, circadian clock regulation plays a key role in optimizing metabolic pathways, increasing fitness, and controlling responses to biotic and abiotic stresses.

In the allotetraploids, the increased level of heterosis is likely caused by interspecific hybridization as well as genome doubling. However, it is unknown how genome dosage and allelic effects influence heterosis, and whether additional clock output traits, such as stress responses, are altered in hybrids. In three related projects, the effects of genomic hybridization (including parent-of-origin effects) and genome dosage on heterosis were elucidated. In my first project, I found that although ploidy influenced many traits, including seed and cell size, biomass and circadian clock gene expression were most strongly influenced by hybridization. Additionally, parent-of-origin effects

between reciprocal hybrids were frequently observed for many traits. In my second project, I described a unique role for RNA-directed DNA methylation (mainly CHH methylation) in mediating the parent-of-origin effect on expression of the circadian clock gene *CCA1* in reciprocal hybrids. Altered *CCA1* expression peaks were associated with heterosis of biomass accumulation in the reciprocal hybrids. Lastly, I used transcriptome sequencing in hybrids at different times of day to examine changes in downstream clock-regulated pathways. In the hybrids, many genes in photosynthetic pathways were upregulated, while many genes involved in biotic and abiotic stresses were repressed during the morning and afternoon, respectively. Additionally, natural variation between parents in stress-responsive gene expression was found to be crucial for producing vigorous hybrids. These conceptual advances increase the mechanistic understanding of heterosis, and may guide selection of parents for making better hybrids.

Table of Contents

List of Tables x

List of Figures xi

Chapter 1. Background and Introduction 1

Central concepts of heterosis 1

Gene expression changes in hybrids 3

Epigenomic changes in hybrids 3

The role of circadian clock-regulated pathways in heterosis 6

Chapter 2. Ploidy and hybridity effects on growth vigor and gene expression in *Arabidopsis thaliana* hybrids and their parents 10

Background and rationale 10

Materials and methods 11

Plant growth and materials 11

Chromosome spreads and flow cytometry 11

DNA and RNA extraction and analysis 12

Biomass measurement 13

Stomatal measurement 13

Seed size and weight analysis 14

Starch and chlorophyll analysis 14

Results 14

Validation of chromosomal content and ploidy levels of F1 hybrids 14

Effects of genome dosage and hybridization on biomass and cell size in F1 hybrids at different ploidy levels 18

Effects of genome dosage and hybridization on expression of circadian clock genes and starch metabolic genes and output traits in hybrids 22

Discussion 27

Chapter 3. A role for CHH methylation in the parent-of-origin effect on altered circadian rhythms and growth vigor in plant hybrids 31
Background and rationale 31
Materials and methods 32
Plant growth and materials....32
Transgenic plants expressing luciferase reporter....33
Embryo dissection and culture (developed by Dr. Helen H. Yu)33
Luciferase assays and data analysis 34
RNA preparation and reverse transcription-polymerase chain reaction (RT-PCR) 35
Genomic DNA extraction and bisulfite sequencing 36
Biomass analysis....37
Results....37
Parent-of-origin effects on biomass heterosis in hybrids....37
Parent-of-origin effects on circadian rhythms 38
CHH methylation and AGO4 affect parent-of-origin effects on circadian gene expression 40
Parent-of-origin effects on CCA1 expression during early stages of embryo development....50
A model for parent-of-origin effects on altered circadian rhythms and growth vigor in hybrids....54
Chapter 4. Natural variation and timing of stress responses promotes heterosis in hybrids....58
Background and rationale 58
Materials and methods 59
Plant growth and materials....59
Biomass measurement 60
Trypan blue staining 60
mRNA-seq library preparation 60
mRNA-seq read mapping and expression quantification 60
Bioinformatics analyses....60
RNA Extraction, cDNA synthesis, and qRT-PCR....61

Plasmid constructs 63
Luciferase assays 64
Regression analysis 65
Cold- and SA-treatments 65
Results 66
Altered expression rhythms of photosynthetic and stress-responsive genes in F1 hybrids 66
Natural variation in stress-responsive gene expression as one predictor for heterosis 73
Circadian clock regulates rhythmic expression of stress-responsive genes 82
Effects of repressing and overexpressing stress-responsive genes on biomass 84
A role for timing of stress-responsive gene induction in biomass heterosis 88
A model for altering stress-responsive gene expression to promote growth vigor in hybrids 93
References 96
Vita……..107

List of Tables

Table 1.1 Primer sequences of *CCA1, LHY, TOC1* and genes involved in photosynthesis and starch metabolism for quantitative RT-PCR 13

Table 3.1 Primers for qRT-PCR and methylation assays 36

List of Figures

Figure 1.1 Altered circadian rhythms in hybrids 8

Figure 2.1 Validation of ploidy and genotype in *Arabidopsis thaliana* hybrids and their parents (Col and C24). 16

Figure 2.2 Validation of genotype and ploidy in ColxC24 and ColXL*er* hybrids and parents. 17

Figure 2.3 Relative effects of genome dosage and hybridization on rosette size, dry weight, stomata size, and stomata density on ploidy hybrids and their parents (Col and C24). 19

Figure 2.4 Seed size, weight, and flower size of ploidy hybrids and their parents.21

Figure 2.5 Expression of the circadian clock genes *CCA1* and *TOC1* at ZT6 and ZT15, and starch and chlorophyll content in ploidy hybrids and their parents. 23

Figure 2.6 Expression of circadian clock genes *CCA1* and *TOC1* in ColXL*er* hybrids and parents at ZT6 and ZT15. 24

Figure 2.7. Expression of the genes involved in chlorophyll and starch metabolism in ColXC24 hybrids and parents at ZT6. 26

Figure 3.1 Parent-of-origin effects on *CCA1* expression and biomass accumulation in hybrids.....39

Figure 3.1 Bisulfite sequencing analysis of DNA methylation at *ASA1*.....42

Figure 3.2 Bisulfite sequencing analysis of DNA methylation in reciprocal hybrids 43

Figure 3.3 Bisulfite sequencing analysis of DNA methylation at the 5'UTR of *CCA1* 44

Figure 3.4 Parent-of-origin effects on *CCA1* expression depend on CHH methylation and AGO4 in reciprocal hybrids 48
Figure 3.5 Biomass analysis in reciprocal hybrids and their parents 49
Figure 3.6 Diurnal expression of clock regulators in developing siliques in *A. thaliana* (Col-0 and L*er*) 51
Figure 3.7 Parent-of-origin effect on circadian rhythms and embryo size in hybrids 52
Figure 3.8 Parent-of-origin effects of *proCCA1:LUC* expression in embryos of reciprocal hybrids 53
Figure 3.9 A model for the parent-of-origin effect on circadian rhythms and growth vigor in hybrids 56
Figure 4.1 Expression of many abiotic and biotic stress-responsive genes is altered in reciprocal F1 hybrids relative to the parents 67
Figure 4.2 GO classifications of DEGs 68
Figure 4.3 Photosynthetic genes are upregulated in F1 hybrids 70
Figure 4.4 Biotic and abiotic stress genes show altered expression in F1 hybrids 71
Figure 4.5 Perturbation of circadian genes in F1 hybrids 72
Figure 4.6 *ACD6* and *COR78* were subjected to diurnal regulation and differentially expressed between two ecotypes, Col and C24 74
Figure 4.7 Natural variation of stress responsive gene expression at specific times of day is associated with biomass heterosis 76
Figure 4.8 Stress-responsive gene expression in vigorous F1 hybrids 77
Figure 4.9 Stress-responsive gene expression in low-vigor F1 hybrids 78
Figure 4.10 Altered expression of *CCA1* and *TOC1* in F1 hybrids with high levels of heterosis 80

Figure 4.11 Altered expression of *CCA1* and *TOC1* in hybrids with low levels of heterosis 81
Figure 4.12 Rhythmic expression patterns of *ACD6* and *COR78* are mediated by components of the circadian clock.....83
Figure 4.13 Effects of knockdown or overexpression of abiotic and biotic genes on biomass 85
Figure 4.14 The induction of stress-responsive gene expression in F1(ColXC24) hybrids and their parents (Col and C24) 86
Figure 4.15 Altered expression of circadian clock genes after stress-treatment....87
Figure 4.16 Effects of cold stress and salicylic acid (SA) on the growth rate and biomass in hybrids and their parents.....90
Figure 4.17 Longer-term cold stress and salicylic acid (SA) affect seedling growth rate and gene expression in F1 (ColXC24) hybrids and their parents91
Figure 4.18 A model for how altered stress-responsive gene expression promotes growth vigor in hybrids.....92
Figure 4.19 Sequence read densities and distributions of DNA methylation, small RNA, and mRNA in the vicinity of *PR1* and *ACD6* genomic regions94

Chapter 1. Background and Introduction

Central concepts of heterosis

Scientific descriptions of hybrid vigor date back to 1876 when Charles Darwin observed that out-crossed plants were more vigorous than self-fertilized plants [1], and in 1908 hybrid vigor was rediscovered and applied to maize breeding by George Shull and Earnest East [2, 3]. Hybrid vigor, also known as heterosis, is a phenomenon in which hybrids between different varieties or species of plants or animals display superior growth and fitness over their parents. Hybridization within and between species is a naturally occurring process and is estimated to take place in at least 25% and 10% of plant and animal species, respectively, and may play an important role in speciation [4, 5]. Because of substantial yield increases in hybrids, many crops and vegetables, including maize and cabbage, are now grown as hybrids [6]. The yield increases in hybrids range from 15-50%, depending on the crop [7].

Arabidopsis thaliana is a useful model to examine the underlying mechanism(s) for heterosis, as intraspecific F1 hybrids between certain ecotypes show increased biomass and seed yield, whereas some ecotype combinations yield non-vigorous hybrids [8, 9]. There is a plethora of naturally occurring ecotypes which differ in many different traits, such as stress tolerance, as well as their geographical locations of origin [10-12]. Since genetic variation between parents is crucial for heterosis [6, 13, 14], the diversity in *A. thaliana* provides a unique opportunity to identify characteristics of parents that produce highly vigorous hybrids. F1 hybrids are genetically unstable, making it difficult to dissect the molecular basis for heterosis. One solution is to develop *Arabidopsis* allotetraploids in which the chromosomes from interspecific hybrids are doubled. As a result, the heterozygosity and hybrid vigor are permanently fixed [4]. Polyploids display

progressive heterosis, with the level of heterosis increasing as the genome diversity in the polyploid increases. Notably, the level of vigor seen in allotetraploids is higher than that in intraspecific hybrids between ecotypes. This observation is consistent with the hypothesis that genetic distance between the parents is critical for heterosis [14, 15].

Although heterosis has been fundamental to crop production since the introduction of hybrid maize in the 1930s [6, 16], the molecular basis still evades elucidation. When heterosis began to be extensively utilized in agriculture, two main hypotheses were proposed to explain heterosis. The original model was the dominance hypothesis, which states that heterosis arises from the complementation of deleterious alleles from one parent with favorable alleles from the other parent [2, 17]. The dominance model predicts that if all deleterious alleles are purged from the parents, the resulting hybrid would no longer show heterosis. This has not been observed, as maize hybrids made from superior parental inbred lines, which were highly selected for yield improvement over many years, still showed the same magnitude of heterosis as hybrids between lesser quality inbreds [18, 19]. The other classical model to explain heterosis is overdominance, which states that different alleles interact in the hybrid leading to improved performance [6, 14]. Although there have been examples of both overdominance and dominance, and rice hybrids have been shown to exhibit all possible models [20, 21], neither can fully explain the phenomenon of heterosis. It has been suggested that these terms are no longer useful for understanding heterosis, as they limit the interpretation of data and were proposed with an insufficient understanding of genetic principles, and that current scientific thinking should move beyond these concepts [6, 14, 22]. The increasing availability of genome-wide approaches will be more useful for developing new insights into the mechanism of heterosis.

Gene expression changes in hybrids

Microarray and RNA-seq approaches have identified genome-wide transcriptome changes, where the expression level of genes deviate from the expected average between the parental values (also known as the mid-parent value, or MPV), in maize hybrids [23-28], rice hybrids [29], and *A. thaliana* hybrids [30, 31]. Other genome-wide gene expression studies have documented nonadditive gene expression in genetically stable *Arabidopsis* allotetraploids derived from *Arabidopsis thaliana* and *Arabidopsis arenosa* [32-34]. Similar gene expression changes have also been found in allopolyploids of *Brassica* [35, 36], cotton [37, 38], *Senecio* [39], *Spartina* [40], *Tragopogon* [41, 42], and wheat [43, 44]. In general, the degree of non-additive expression is higher in interspecific hybrids than in the intraspecific hybrids (i.e. between ecotypes), which is correlated with the higher amount of vigor observed in the interspecific crosses. Although these studies identify the importance of differentially expressed genes (DEGs), little consensus about the types of genes or pathways involved in heterosis has emerged. This is most likely due to variability in species, tissues, developmental stage, sequencing or microarray platforms, and/or time of day of sample collection.

Epigenomic changes in hybrids

In addition to transcriptome changes in hybrids, there is also substantial evidence of large scale proteomic [45-48], metabolomic [49-51], and epigenomic [13, 29, 31, 52] upheaval. Although proteomic and metabolomic analysis have proven useful for the prediction of heterosis, epigenomic changes will be further discussed here.

It is predicted that epigenetic variation between the parents is crucial for heterosis [6, 14]. This is especially true for intraspecific hybrids generated from genetically similar parents [53]. For example, *A. thaliana* ecotypes show epigenetic diversity in levels of 24-nucleotide small interfering RNAs (siRNAs) [13], chromatin modifications [54], and

DNA methylation [31, 52]. This has also been observed in rice subspecies [29] and maize inbreds [55]. It is plausible that hybrids between parents with extensive epigenetic diversity are more vigorous than hybrids from more epigenetically similar parents, and this hypothesis is consistent with the observation that more vigorous hybrids show greater levels of differential gene expression than less vigorous hybrids [30], possibly as a result of the epigenetic diversity in the parents. Regardless, current evidence strongly suggests that epigenetic divergence between the parents is necessary for subsequent epigenetic changes in the hybrids, which occur most frequently at loci that are differentially regulated between the parents [13, 31, 53, 56].

Recent studies document DNA methylation level changes in hybrids of *Arabidopsis* [13, 31, 52] and rice [29, 57]. In *A. thaliana* hybrids between ecotypes L*er* and C24, methylation levels in all sequence contexts (CG, CHG, and CHH, where H≠ G) were increased, especially in regions containing transposable elements (TEs). The majority of regions with higher methylation also generated siRNAs and were differentially methylated between the parents [31]. When DNA methylation was chemically removed using 5-aza-2'-deoxycytidine, growth vigor was compromised in hybrids [31]. Other studies in hybrids between L*er* and C24 found that methylation changes in hybrids are most frequently due to Trans Chromosomal Methylation (TCM) and Trans Chromosomal deMethylation (TCdM), where the methylation level of one parental allele is altered to resemble that of the other parent [52]. TCM and TCdM are frequently associated with siRNAs, but there are some loci where TCM and TCdM occur in regions with no siRNAs, indicating that chromatin modifications may also be important for methylation changes [56]. The TCM and TCdM events persist in F2 plants, and the segregation of epigenetic determinants in F2 populations of hybrids may lead to the reduction in vigor seen after the F1 generation [56].

Small RNAs (sRNAs), including both siRNAs and microRNAs (miRNAs), are associated with heterosis in *Arabidopsis* [13, 31, 58, 59], rice [29, 57], wheat [60], and maize [61]. Although the majority of 24 nt siRNAs are expressed at mid-parent levels, some studies found downregulation of a small percentage of loci [13, 29]. Another study found that 21 nt sRNAs (primarily miRNAs) increased in hybrids, whereas 24 nt siRNAs remained largely unchanged [31]. The authors propose that this is due to varying levels in sRNA sequence coverage between the studies, and upon closer examination of previous data it was found that rice hybrids did not display decreased levels of siRNAs [29, 31]. Regardless of these discrepancies, it is clear that sRNAs show expression changes in hybrids, although the role they play in promoting heterosis is unclear. Shen *et al.* generated hybrids lacking *HEN1*, a crucial gene responsible for all sRNA (including both siRNAs and miRNAs) biogenesis, and found that heterosis was significantly reduced [31], whereas maize hybrids unable to generate 24 nt siRNAs maintained wild-type levels of vigor [61]. This disagreement between the two studies may be caused by the difference in the affected sRNA populations in the hybrids.

Histone modifications may also be another epigenetic component contributing to heterosis. Histones can be acetylated on lysine residues (e.g. H3K9ac), typically associated with gene activation, or methylated on lysine residues, which is associated with both gene activation (e.g. H3K4me3) and repression (e.g. H3K27me3) [14, 62]. In *Arabidopsis* [54] and rice [29], inheritance of histone modifications in hybrids is largely additive. However, when histone modifications were inherited non-additively, transcript changes were also observed in rice hybrids [29]. In *Arabidopsis* allopolyploids, altered expression of circadian clock regulators *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*) was associated with non-additive inheritance of both H3K9ac and H3K4me2 in the promoter regions of these genes [63].

The contribution of the circadian clock and associated pathways will be further discussed in the next section.

The role of circadian clock-regulated pathways in heterosis

Genome-wide evidence suggests that heterosis arises from allelic interactions between parental genomes, leading to altered programming of gene networks that promote growth, stress tolerance, and fitness of hybrids [14]. *A. thaliana* hybrids were shown to have large scale gene expression changes in photosynthesis and chlorophyll biosynthesis genes, which contributes to an increased photosynthetic capacity in the hybrids compared to the parents [30]. These results are consistent with previous studies which found extensive nonadditive expression of genes involved in energy production and metabolism in *Arabidopsis* allotetraploids [32, 63]. There is mounting evidence that stress-responsive pathways are also changed in hybrids and allotetraploids, as many abiotic and biotic stress-responsive genes are repressed in *Arabidopsis* allotetraploids [32, 64-66]. Many of the pathways that are differentially expressed in hybrids are under the control of the circadian clock and are considered output pathways of the clock.

Most living organisms have adapted to 24-hour day-night cycles. In *Arabidopsis*, approximately one third of expressed genes are clock controlled [67], as well as most aspects of plant growth and development including coordination of processes with daily rhythms and seasons [68]. Circadian-mediated regulatory networks, which promote growth and fitness in plants and animals, control output pathways associated with heterosis [69-71]. In humans, disruption of circadian rhythms leads to health problems, including obesity, type 2 diabetes, and cardiovascular diseases [71, 72]. Mice without a functional circadian clock lack a diurnal feeding rhythm, are obese, and develop metabolic syndromes [73].

In *Arabidopsis* allotetraploids and *A. thaliana* F1 hybrids, epigenetic regulation of circadian clock genes, *CCA1* and *LHY* and their reciprocal regulators *TIMING OF CAB EXPRESSION1* (*TOC1*) and *GIGANTEA* (*GI*), leads to altered circadian rhythms, which in turn promotes expression of photosynthetic and metabolic pathways [63]. Consequently, chlorophyll, starch, and sugar contents in vegetative tissues are increased during the day, so more can be utilized and degraded at night to promote growth [74]. In addition, altering expression amplitudes of circadian clock genes using RNA interference (RNAi) also increases or decreases starch amount and rosette size [63]. This correlation of circadian regulation with heterosis has also been observed in additional hybrids of *A. thaliana* [31, 75]. In super-hybrid rice, expression changes in circadian clock genes and their regulatory networks are also related with yield quantitative trait loci (QTLs) [76].

In *A. thaliana* diploids, the circadian clock also regulates many other biological processes including responses to abiotic [77, 78] and biotic [79-82] stresses, and plant hormones such as auxin, cytokinins, abscisic acid (ABA), jasmonic acid (JA), and salicylic acid (SA) [78, 80, 83-85]. Cold-inducible genes have diurnal expression patterns and are enriched in evening elements that bind circadian and other transcription factors, thus integrating circadian regulation with cold-inducible gene expression [77]. Additionally, plant sensitivity to cold is gated by the circadian clock, with time of day of cold exposure influencing levels of induction of cold-responsive genes [86-90]. The clock also mediates responses to biotic stress and synchronizes defense responses with insect circadian behavior [80, 91]. Clock regulators modulate defense genes that respond to powdery mildew infection, allowing plants to “anticipate” infection at dawn and to time immune responses after infection [79]. Animal immune responses may also be influenced by the circadian clock [92].

Since *Arabidopsis* allotetraploids and *A. thaliana* F1 hybrids have altered expression of circadian clock genes, it is plausible that clock-regulated stress genes and other output pathways may also be affected. Consistent with this hypothesis, many stress-responsive genes, including abiotic and biotic stress-responsive genes, are generally repressed in *Arabidopsis* allotetraploids [32], suggesting a correlation between heterosis with repression of stress-responsive genes. Moreover, increased expression of biotic and abiotic stress-responsive genes often reduces fitness and growth [93-97]. Knockdown of *ACCELERATED CELL DEATH 6* (*ACD6*), a broad-spectrum disease resistance gene, increases biomass accumulation in certain ecotypes [98]. *A. thaliana* F1 hybrids are more cold tolerant, yet also accumulate more biomass than the parents [99, 100]. Photosynthesis is one of the processes most strongly affected by both abiotic and biotic stresses [101, 102], and *A. thaliana* hybrids have been shown to have an increased capacity for photosynthesis [30]. These data suggest that hybrids may balance stress responses and energy production via the circadian clock to enhance growth and productivity (Figure 1.1).

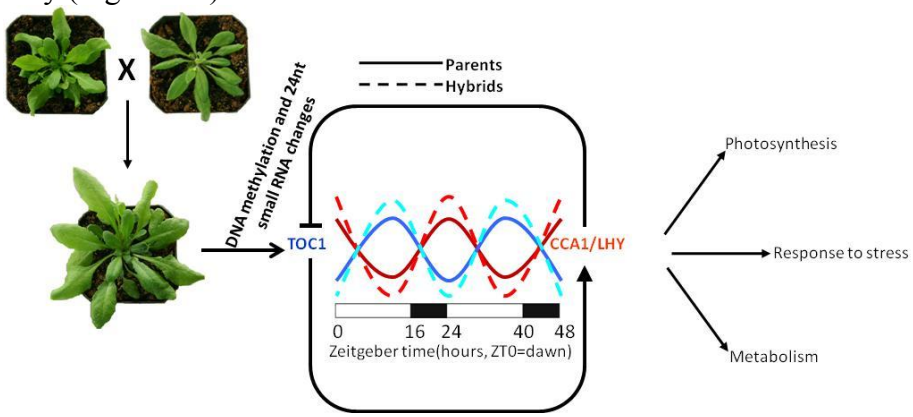


Figure 1.1 Altered circadian rhythms in hybrids

In the hybrids, the genomic interactions between parents induce epigenetic changes in *CCA1* and *LHY*, leading to changes in expression amplitudes (red and blue dashed lines) relative to the expression values in the parents (solid red and blue lines). Altered expression of *CCA1* in the hybrids leads to expression changes of circadian clock-output genes in various output pathways, including photosynthesis, stress responses, and metabolism.

In this dissertation, the roles of polyploidy, circadian rhythms, and stress responses in hybrid vigor were investigated. Since the cause of growth vigor in allopolyploids is confounded by the effects of increased levels of ploidy and genomic hybridity, the relative effects of polyploidy and hybridization were systematically investigated using *A. thaliana* F1 intraspecific diploid, triploid, and tetraploid hybrids. In addition, the role of the circadian clock in mediating parent-of-origin effects on heterosis was examined in reciprocal hybrids. To identify other circadian-regulated genes involved in heterosis, transcriptome sequencing at three different times of day was used to analyze global gene expression in F1 intraspecific hybrids. A deeper understanding of these different aspects of growth vigor in *Arabidopsis* hybrids and allopolyploids will allow for developing cost-effective methods to increase biomass and yields for food and energy crops, many of which are grown as hybrids and polyploids.

Chapter 2. Ploidy and hybridity effects on growth vigor and gene expression in *Arabidopsis thaliana* hybrids and their parents

Background and rationale

After the F1 generation, genetic determinants controlling heterosis segregate, leading to difficulties when studying the molecular basis for heterosis. Allotetraploids, in which the chromosomes in interspecific hybrids are doubled, are genetically stable. As a result, the heterozygosity and hybrid vigor are permanently fixed in allopolyploids [4, 22]. Growth vigor in allopolyploids is confounded by the effects of increased levels of ploidy and genomic hybridity. There is no obvious growth vigor in *A. thaliana* autotetraploids, although genome dosage is increased relative to diploids [4]. In maize, the number of nonadditively expressed genes and the degree of their expression increase from duplex to quadruplex hybrids [103, 104]. These gene expression trends are consistent with observed phenotypic variation, but the mechanistic connections remain unknown. To test the relative impact of ploidy and hybridity on heterosis, a series of reciprocal diploid, triploid, and tetraploid hybrids were generated using *A. thaliana* ecotypes (Col, C24, and Ler) of diploid and isogenic tetraploid parents. First, these genetic materials were validated using chromosome counts, flow cytometry, and DNA markers. Biomass, stomatal size, flower morphology, and seed size and weight were then evaluated in these lines. Finally, expression of circadian clock genes was examined to determine if there is a correlation with starch and chlorophyll contents and biomass in these hybrids. The results provide novel insights into the effects of genomic hybridization, ploidy, and circadian gene expression on biomass, cell and seed size, and starch metabolism in *Arabidopsis* intraspecific hybrids. The information is of practical use for the improvement of biomass and seed production in vegetable and food crops.

Materials and methods

Plant growth and materials

Two sets of reciprocal hybrids were generated in each genotypic combination using *A. thaliana* ecotypes Columbia (Col), C24, and Landsberg *erecta* (L*er*) diploid (2X) and isogenic autotetraploids (4X) as parents (Col2, Col4, L*er*2, L*er*4, C24-2, and C24-4). These hybrids include (1) reciprocal diploid hybrids (Col2XC24-2 and C24-2XCol2; by convention the maternal parent is listed prior to the paternal parent in a genetic cross, and Col2XL*er*2 and L*er*2XCol2), reciprocal triploid hybrids (Col2XC24-4 and C24-4XCol2; Col4XC24-2 and C24-2XCol4; and Col2XL*er*4), and reciprocal tetraploid hybrids (Col4XC24-4 and C24-4XCol4; Col4XL*er*4 and L*er*4XCol4). Crossing within (parents) and between *A. thaliana* ecotypes was carried out by removing immature anthers from unopened flower buds in the maternal ecotype and fertilizing the stigma with pollen from freshly opened flowers of the paternal ecotype. Plants were grown on soil in 16/8-h (light/dark) cycles [33].

Chromosome spreads and flow cytometry

Chromosome spreads of young flower buds were prepared as previously described [105] and stained with 4',6-diamidino-2-phenylindole (DAPI). Slides were observed using a light microscope (Zeiss Axiovert 200 M).

For flow cytometry samples, 70 mg leaves from seedlings were collected and processed as previously described [106], except that the samples were filtered through 30-μm Partec CellTrics® filters and were stained using propidium iodide (PI) at a concentration of 100 μg/ml. Samples were stained for 40 to 60 minutes before performing flow cytometry. Flow cytometry was performed on a BD Biosciences FACSCalibur flow cytometer (BD Biosciences, San Jose, California), and data were acquired using

CellQuestPro software (Becton, Dickinson and Company, Franklin Lakes, New Jersey). Samples were run on low pressure long enough to acquire clear peaks (5-10 minutes). Because PI (emission maximum 639 nm) was used as the DNA stain, the FL2 detector (564 – 606 nm) was used to measure fluorescence. FL2-Area (FL2-A), a measurement of integrated fluorescence signal, was used as the parameter linearly correlated to DNA content [107]. *A. thaliana* diploid leaf tissues were used to calibrate the instrument prior to running the samples. Noise signals derived from subcellular debris were eliminated by gating [66].

DNA and RNA extraction and analysis

Total genomic DNA was isolated from leaves by grinding in extraction buffer (0.2M Tris-HCl pH 7.5, 0.25M NaCl, 25mM EDTA, and 0.5% SDS). The supernatant was then precipitated with isopropanol, and the pellet was rinsed with 75% EtOH. Following centrifugation, the pellet was resuspended in sterile water. PCR was used to amplify DNA fragments with length polymorphisms between different ecotypes and in the hybrids. The marker used to distinguish ecotypes, nga106, is described in [108]. PCR primers were obtained from TAIR. Approximately 500 ng genomic DNA was used for PCR using AmpliTaq DNA Polymerase (Applied Biosystems, Foster City, California).

Tissue collected for gene expression analysis was collected from plants before bolting (6-8 rosette leaves) at indicated Zeitgeber time (ZT0 = dawn) [68]. Total RNA was extracted using Concert Plant RNA Reagent (Invitrogen, Carlsbad, California) according to the manufacturer's instructions. Total RNA was digested with RQ1 RNase-Free DNase (Promega, Madison, Wisconsin) to remove any DNA. cDNA was synthesized from the DNase-treated RNA using Superscript III (Invitrogen). 1 μl of diluted cDNA was used for quantitative RT–PCR analysis using the primer pairs listed in

Table 1 in an ABI7500 machine (Applied Biosystems) as described in [109], except that *ACT* was used as an internal control to calculate the relative expression levels in three biological replications.

Table 1.1 Primer sequences of *CCA1, LHY, TOC1* and genes involved in photosynthesis and starch metabolism for quantitative RT-PCR

LOCUS	NAME	FORWARD PRIMER	REVERSE PRIMER
At2g46830	*CCA1*	5'-CCTCGTCAGACACAGACTTCCA-3'	5'-CCGCAGTAGAATCAGCTCCAATA-3'
At5g61380	*TOC1*	5'-GTTGATGGATCGGGTTTCTC-3'	5'-TCATGACCCCATGCATACAG -3'
At5g09810	*ACT*	5'-GTCTGTGACAATGGAACTGGAA-3'	5'-CTTTCTGACCCATACCAACCAT-3'
At5g54190	*PORA*	5'-GTGGTTGTCACGGGAGCTTC-3'	5'-TGCCTTTGCCGTTGCTAAAC-3'
At4g27440	*PORB*	5'-GTGGACGGCAAGAAAACGTT-3'	5'-GGCTCCAGTGACCACCACAT-3'
At1g69830	*AMY3*	5'-CTTCAAGTAGCTCGCCCGTT-3'	5'-TGGGTTTACTCACTTGGGCAG-3'
At5g64860	*DPE1*	5'-GTTCCGGATCCAGAGAGCAG-3'	5'-CGTCGGGTGTAGCAAAACG-3'
At5g26570	*GWD3*	5'-TTCGCCGGACTTATCATTCG-3'	5'-TCCGGATCAGCTGGACTCAC-3'

Biomass measurement

Whole rosettes from hybrids and parents were harvested at approximately 3 weeks of age (before bolting) and placed in Lawson #217 hybridization bags (Lawson Bags, Northfield, Illinois). Dry weights from rosette leaves were determined after drying the plants at 80°C for 24 h. Each rosette in three biological pools was weighed individually, and the average was used to calculate standard deviations.

Stomatal measurement

Stomatal imprints were taken from the abaxial side of leaves from plants before bolting. Young leaves were placed onto a small drop of cyanoacrylate adhesive ("Superglue") and gentle pressure was applied for approximately 30 seconds. Leaves

were then gently removed, and the imprint was allowed to dry for at least 10 minutes before viewing with a Leica DM LB microscope. Stomatal size was determined by measuring 20 stomata per genotype lengthwise. Stomatal density was defined as the number of stomata per square mm.

Seed size and weight analysis

Average seed weight was determined by weighing mature dry seeds in batches of 150. The weights of three batches were measured for each seed lot using an analytical balance. Sizes of parent and hybrid seeds were analyzed by separating batches of 150 seeds using a series of fine wire sieves. Sieve mesh sizes 40, 45, 50, 60, 70, and 80 (Fisher Scientific, Waltham, Massachusetts) with exclusion sizes of 425, 355, 300, 250, 212, and 180 μm, respectively, were used for each analysis. Seeds retained by each sieve were counted and a weighted average was calculated. Three batches of seeds were measured for each genotype.

Starch and chlorophyll analysis

Starch and chlorophyll extraction and analyses were performed as in [63], except that 300 mg tissue was used per biological replication.

Results

Validation of chromosomal content and ploidy levels of F1 hybrids

Three different methods were employed to ensure that the F1 hybrids used for this study (see Materials and Methods) were true hybrids with the expected ploidy levels (Figure 2.1A-L). To determine the ploidy level of the hybrids, chromosome spreads were prepared from young flower buds and stained with DAPI. All nuclei observed had the expected number of chromosomes. For example, all triploid hybrids contained 15

chromosomes (Figure 2.1E-H), and all tetraploids and tetraploid hybrids had 20 chromosomes (Figure 2.1I-L). Flow cytometry was also used to examine the genome composition (Figures 2.1M-R and 2.2A-F). Consistent with previous findings [110], endoreduplication is commonly observed in leaves of *A. thaliana* diploids (Figures 2.1M and 2.2A). The level of endoploidy is proportionally increased from diploids to tetraploids (Figures 2.1N and 2.2B), consistent with data in a recent study [66]. In addition, no significant difference in the level of endoreduplication was observed in the comparison between hybrids and parents at the same ploidy levels (e.g., diploid parents and hybrids, Figure 2.1M,O). As expected, the endoploidy level in the triploid hybrids was between those in the diploids and tetraploids (Figures 2.1Q,R and 2.2E,F). Genomic changes in different ploidy levels of *A. thaliana* are proportional to their endoploidy levels, suggesting that molecular changes in *A. thaliana* polyploids are not compromised by endoreduplication. However, neither chromosome counts nor flow cytometry in hybrids can rule out a possibility of selfing. Thus, the genomic content of diploid and tetraploid hybrids was further validated using simple sequence length polymorphisms (SSLPs) between Col and C24 or Col and L*er* ecotypes (Figure 2.2G,H). All hybrids genotyped showed presence of both parental fragments.

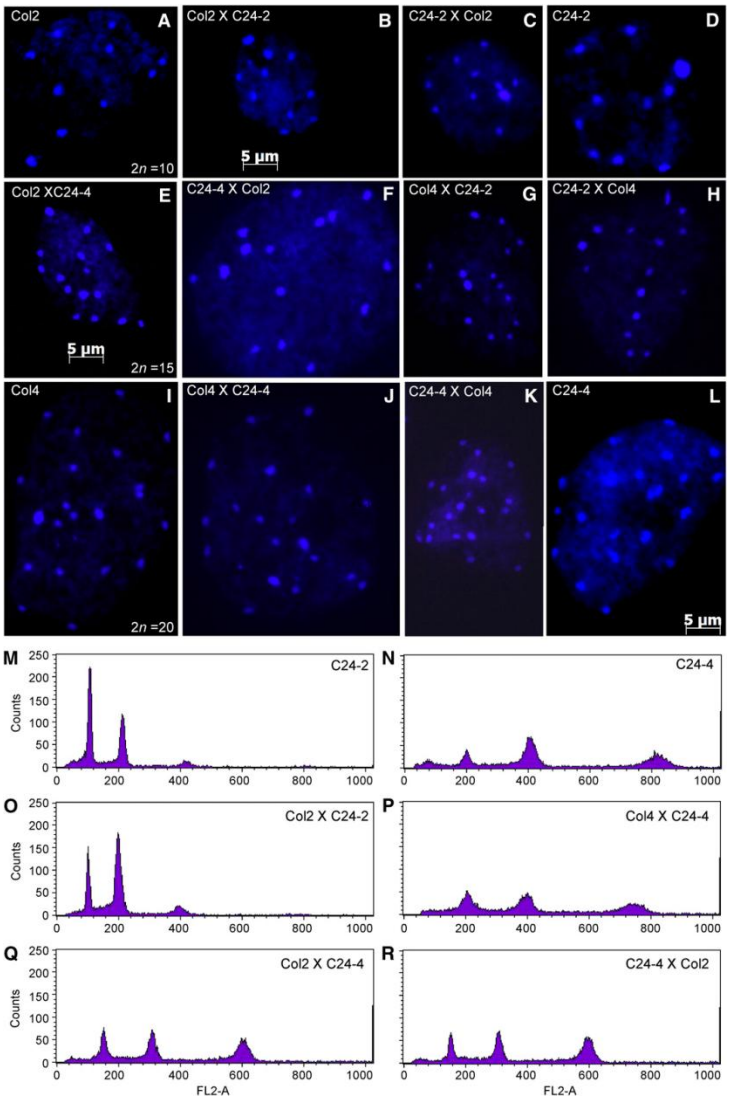


Figure 2.1 Validation of ploidy and genotype in *Arabidopsis thaliana* hybrids and their parents (Col and C24).

(**A–L**) Chromosome spreads were prepared from ploidy hybrids and their parents and stained with DAPI. (**A**) Col2, (**B**) Col2XC24-2, (**C**) C24-2XCol2, (**D**) C24-2, (**E**) Col2XC24-4, (**F**) C24-4XCol2, (**G**) Col4XC24-2, (**H**) C24-2XCol4, (**I**) Col4, (**J**) Col4XC24-4, (**K**) C24-4XCol4, and (**L**) C24-4. One representative nucleus per genotype is shown (scale bar = 5 mm). (**M–R**) Flow cytometry analysis of nuclei from leaves of hybrids and parents. (**M**) C24-2, (**N**) C24-4, (**O**) Col2XC24-2, (**P**) Col4XC24-4, (**Q**) Col2XC24-4, and (**R**) C24-4XCol2. Filtered nuclei were stained with propidium iodide and analyzed using flow cytometry (X-axis = fluorescence intensity, Y axis = nuclei counts).

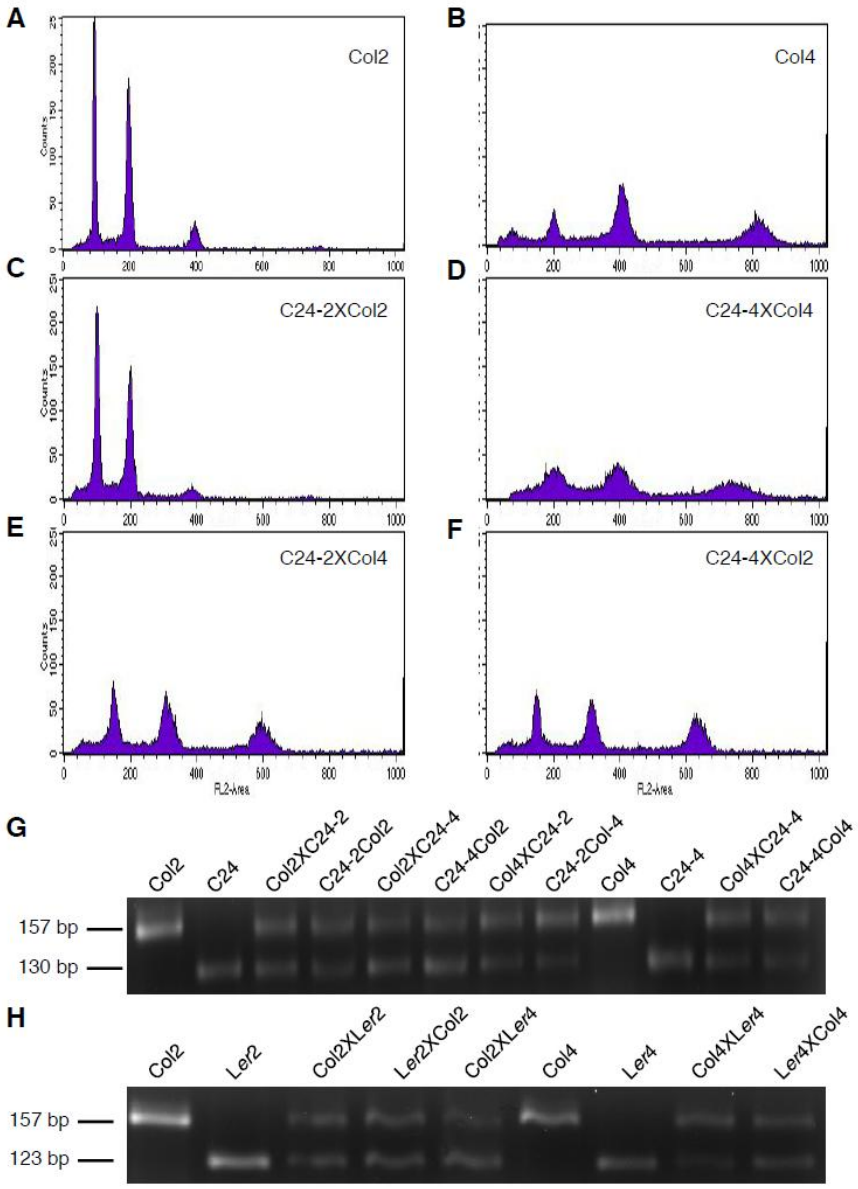


Figure 2.2 Validation of genotype and ploidy in ColxC24 and ColXL*er* hybrids and parents.

(**A-F**) Flow cytometry analysis of nuclei from leaves of hybrids and parents. Filtered nuclei were stained with propidium iodide and analyzed using flow cytometry (X-axis = fluorescence intensity, Y axis = nuclei counts) (**G**) Genotyping in ColXC24 hybrids and parents using genomic DNA PCR. (**H**) Genotyping in ColXL*er* hybrids and parents using genomic DNA PCR. PCR length polymorphisms between the two different ecotypes were resolved on a 4% agarose gel to distinguish the two different genomes present in the hybrids.

Effects of genome dosage and hybridization on biomass and cell size in F1 hybrids at different ploidy levels

Given that some hybrids formed between *A. thaliana* ecotypes do not display obvious growth vigor, reciprocal hybrids of different ploidy levels were generated between Col and C24 because the diploid hybrids are shown to display growth vigor [8, 9]. The diploid hybrids both displayed increased levels of biomass when compared with their parents (Figures 2.3A,B). The biomass between reciprocal hybrids was also significantly different but to a lesser degree (Figure 2.3B). Interestingly, the triploid hybrids showed a much greater size disparity between reciprocal crosses (Figure 2.3A,B). In particular, both triploid hybrids with a tetraploid father are 1-1.5 times larger than the triploid hybrids with a tetraploid mother. This obvious parent-of-origin effect on organismal growth has also been shown in triploid humans [111]. The size difference between reciprocal triploid hybrids became less dramatic after flowering. The triploid hybrids with a tetraploid mother had similar biomass relative to the diploid and tetraploid parents. Somewhat unexpectedly, tetraploid hybrids were slightly larger than the tetraploid parents but not significantly larger than the diploid hybrids. In the reciprocal tetraploid hybrids, the biomass was slightly different (Figure 2.3A, B).

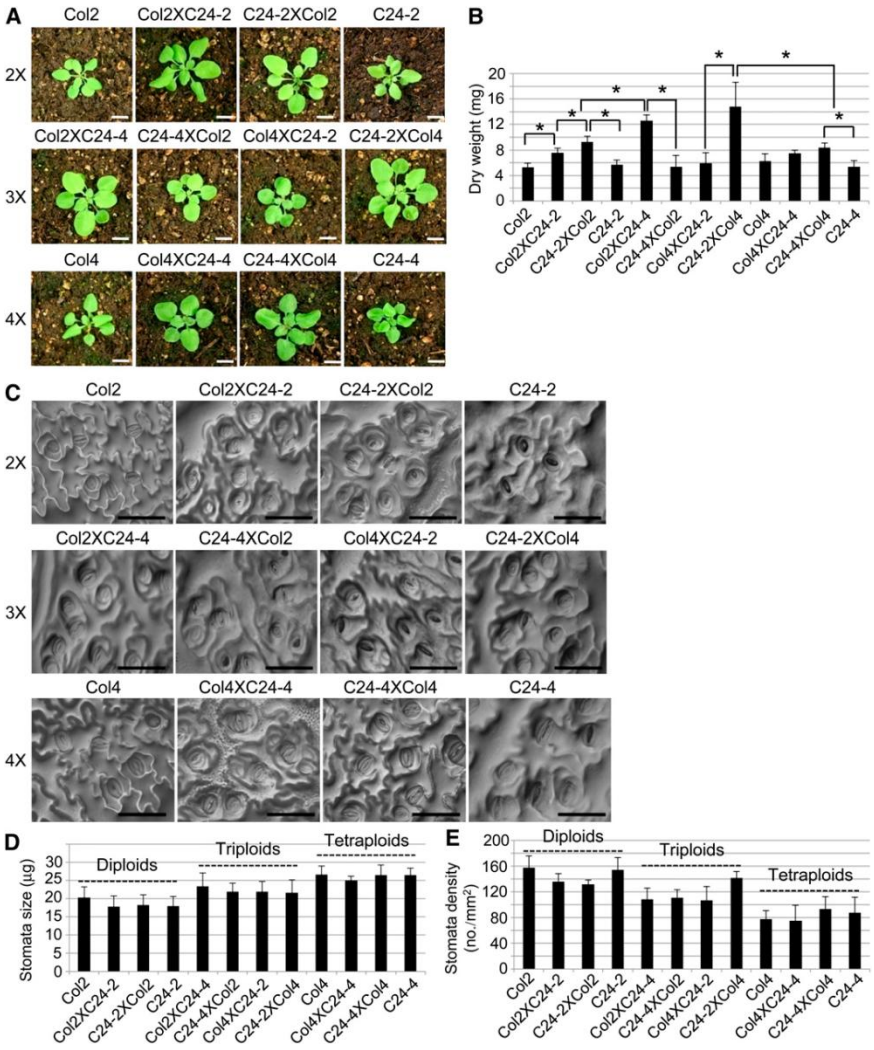


Figure 2.3 Relative effects of genome dosage and hybridization on rosette size, dry weight, stomata size, and stomata density on ploidy hybrids and their parents (Col and C24).

(**A**) Morphological vigor in diploid, triploid, and tetraploid hybrids and their parents. Images are of 15- day-old plants (bar = 1 cm). (**B**) Aerial biomass in diploids, triploid, and tetraploid hybrids and parents (n = 4–5 plants). Asterisks indicate statistical significance (p = 0.05). (**C**) Stomatal imprints of diploid, triploid, and tetraploid hybrids and their parents (bar = 50 mm) (**D**) Stomatal size (n = 20). (**E**) Stomatal density per mm^2 (n = 150–200) in diploid, triploid, and tetraploid hybrids and their parents. Error bar ± SD.

Increased biomass is probably due to increased cell size, number, or both. To test this, stomatal size and density were measured in these lines. Stomata are epidermal structures that consist of two guard cells surrounding a pore that allows for gas exchange [112]. To determine the effects of hybridization and genome dosage on stomatal size and density, stomatal imprints were taken of diploid, triploid, and tetraploid hybrids between Col and C24 ecotypes and quantified (Figure 2.3C-E). The stomatal size increased from diploids, triploids, to tetraploids, whereas the density was negatively correlated with the genome dosage. No obvious difference was found between hybrids and parents at different ploidy levels, neither was there a difference between the reciprocal hybrids, with a couple of exceptions. The triploid hybrid (C24-2XCol4) showed higher stomatal density than other triploid hybrids. Both Col and C24 diploids had higher cell density than their hybrids, suggesting larger pavement cells in hybrids than in the parents.

Although ploidy did not obviously affect biomass, seed size and weight as well as flower size are more strongly affected by the ploidy levels. In general, increasing the ploidy level is positively correlated with increasing the flower size and seed weight so tetraploids had heavier seeds and larger flowers than diploids (Figure 2.4A-D). The difference in seed size in some comparisons was not dramatic probably because of a relatively crude measurement method. Interestingly, seeds in one diploid hybrid (C24-2XCol2) were nearly as large as tetraploid hybrid seeds and 20-40% larger and 60-120% heavier than their respective parents (Figure 2.4C,D). There was a significant parent-of-origin effect on seed weight in reciprocal diploid hybrids, although this effect on seeds was not obvious in the reciprocal tetraploid hybrids. The parent-of-origin effect was most obvious in triploid hybrid seeds. Similar to vegetative growth and biomass, the triploid hybrids with a tetraploid father produced larger and heavier seeds than the triploid

hybrids with a tetraploid mother. Flower size in triploid hybrids also showed this same trend. This obvious parent-of-origin effect is interesting and remains to be tested.

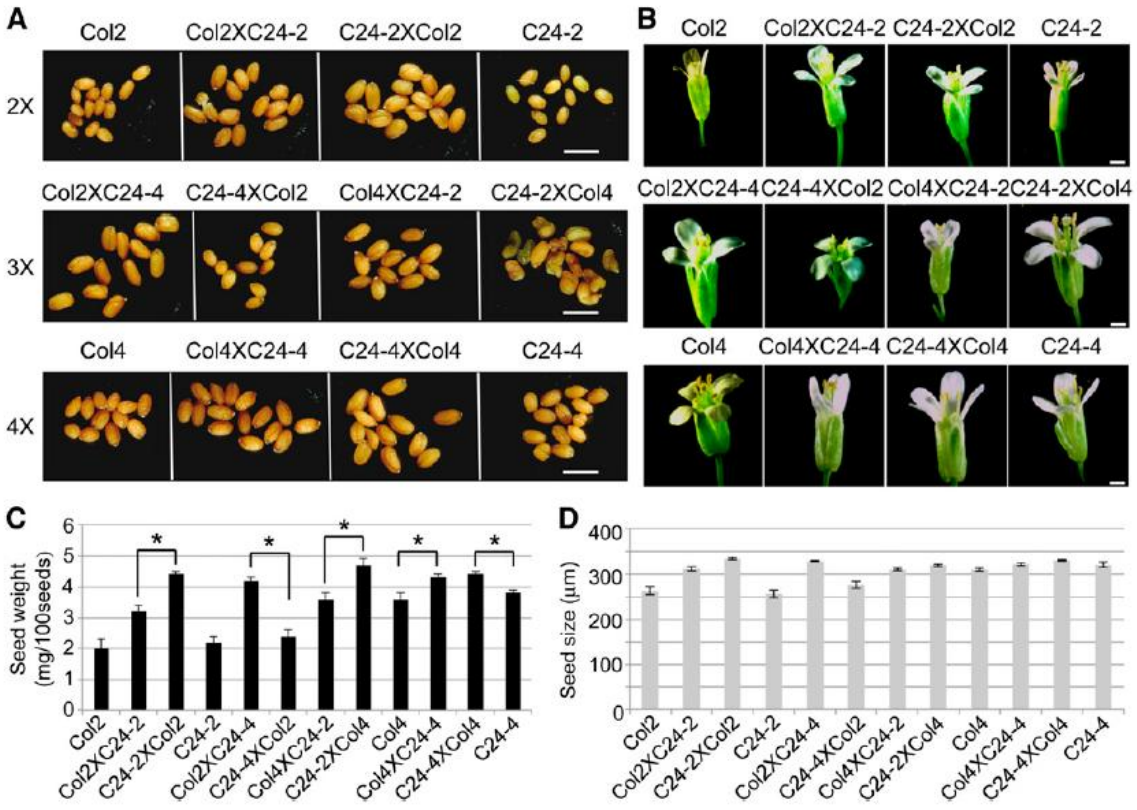


Figure 2.4 Seed size, weight, and flower size of ploidy hybrids and their parents.

(**A**) Seeds of diploid, triploid, and tetraploid hybrids and their parents (bar =1 mm). (**B**) Flower size in diploid, triploid, and tetraploid hybrids and their parents (bar = 1 mm). (**C**) Seed weight per 100 seeds (n = 150). Asterisks indicate statistical significance (p = 0.05). (**D**) Seed size measured using mesh sieves (n = 150). Error bar ± SD.

Effects of genome dosage and hybridization on expression of circadian clock genes and starch metabolic genes and output traits in hybrids

A recent study showed that in *Arabidopsis* hybrids and allopolyploids epigenetic regulation of core clock genes is directly linked to increased expression of downstream genes and metabolic pathways, leading to increased amounts of chlorophylls, starch, and sugars during vegetative growth [63]. Expression changes in circadian clock genes and their regulatory networks are also related with yield QTLs in super-hybrid rice [76]. In *Arabidopsis*, the core loop of the circadian clock is composed of the transcription factors *CCA1* and *LHY*, which are positively regulated by *TOC1* and *CHE* [68, 113, 114]. *CCA1* and *LHY* are negative regulators of *TOC1*, and *TOC1* and *CHE* positively regulate the expression of *CCA1* and *LHY*. To determine the effects of ploidy and hybridization on gene expression and circadian-mediated growth vigor, expression of core clock regulators and their downstream genes in hybrids with different ploidy levels was examined. Consistent with the published data [63], in diploid hybrids between Col and C24 *CCA1* transcript levels were reduced 1.5-2 fold relative to the mid-parent value (MPV) at ZT6 and slightly upregulated at ZT15 (Figure 2.5A). Conversely, *TOC1* was upregulated at ZT6 and downregulated at ZT15 relative to MPV (~2.5 fold) (Figure 2.5B). A similar trend of *CCA1* repression and *TOC1* upregulation in the triploid and tetraploid hybrids was also found. *CCA1* was repressed 2-4 fold in the triploid hybrids and ~1.5-2 fold in the tetraploid hybrids at ZT6 and upregulated at ZT15. *TOC1* was upregulated at ZT6 and downregulated at ZT15 in the triploid hybrids. Interestingly, the levels of *CCA1* repression and *TOC1* upregulation were different between reciprocal hybrids in the diploid, triploid and tetraploid levels, respectively, which are generally correlated with parent-of-origin effects on biomass (Figure 2.3A,B). However, some changes were not

obvious between ploidy levels. The molecular connection between changes in clock genes and biomass in reciprocal crosses needs to be further investigated.

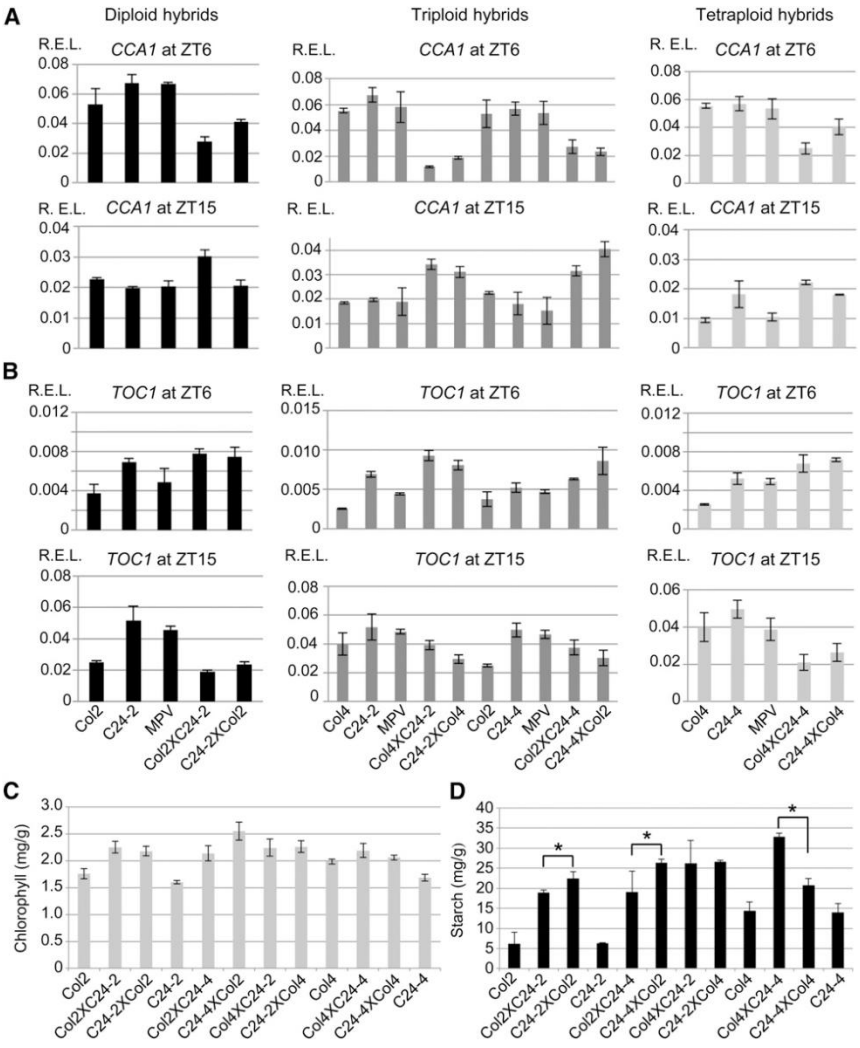


Figure 2.5 Expression of the circadian clock genes *CCA1* and *TOC1* at ZT6 and ZT15, and starch and chlorophyll content in ploidy hybrids and their parents.

Quantitative RT-PCR analysis of (**A**) *CCA1* and (**B**) *TOC1* (n = 3, *ACT* as an internal control) in diploid, triploid, and tetraploid hybrids and their parents. Error bar ± SD. The (**C**) chlorophyll content (n = 2) and (**D**) starch content (n = 3) in diploid, triploid, and tetraploid hybrids and their parents. Asterisks indicate statistical significance (p = 0.05). Error bar ± SD.

Consistent with their similar rosette size, parental lines of diploid and tetraploid Col or C24 showed similar expression levels of *CCA1* and *TOC1* (Figure 2.5A,B), which provides another piece of correlative data between circadian rhythms and growth vigor [63]. Hybrids between Col and L*er* also showed a general trend of *CCA1* repression at ZT6 and upregulation at ZT15, but to a lesser extent (Figure 2.6). This is probably related to the level of hybrid vigor that is dependent on genotypic combinations, as ColXL*er* is a less vigorous hybrid [9, 15]. In general, the larger the genetic distance between the parents, the higher the level of heterosis, if hybrids can be formed [4].

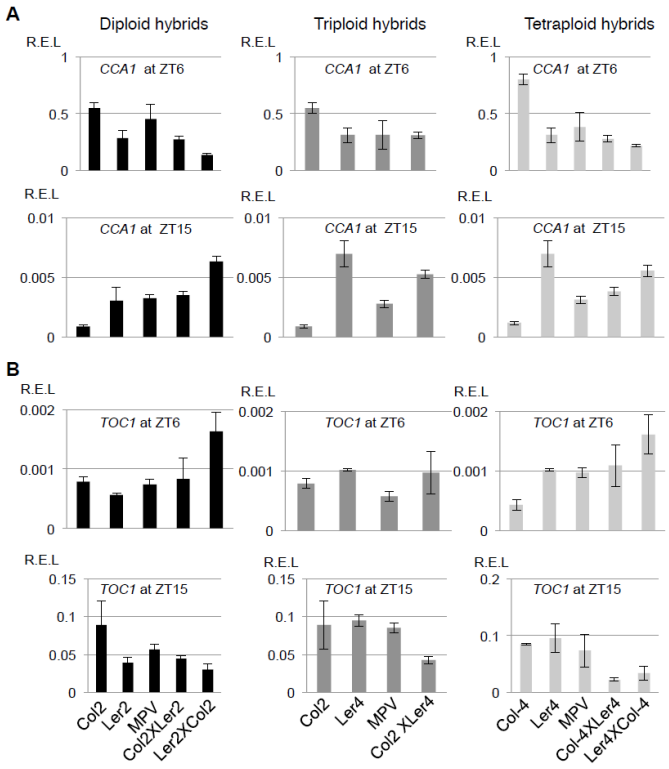


Figure 2.6 Expression of circadian clock genes *CCA1* and *TOC1* in ColXL*er* hybrids and parents at ZT6 and ZT15.

(**A**) Quantitative RT-PCR analysis of *CCA1* and (**B**) *TOC1* (n=3, *ACT* as an internal control). Error bars ± SD.

The expression of downstream genes and circadian-mediated output traits, including starch and chlorophyll content, was examined in diploid, triploid, and tetraploid hybrids. *PORA* and *PORB* catalyze the only known light-requiring step of tetrapyrrole biosynthesis [115]. Both *PORA* and *PORB* contain evening elements (EE) and *CCA1* binding sites (CBS) within their promoters, and are the targets of *CCA1* [86]. Both *PORA* and *PORB* were upregulated in the diploid and triploid hybrids between Col and C24 (Figure 2.7), but were less induced in the tetraploid hybrids than in the diploid and triploid hybrids. As a result, total chlorophyll content was increased in all hybrids relative to their respective parents, with a smaller level of increase in the tetraploid hybrids (Figure 2.5C).

In *Arabidopsis* allopolyploids, many EE and CBS-containing genes involved in starch metabolism are also upregulated [63]. *DPE1*, *AMY3*, and *GWD3* were upregulated in the hybrids between Col and C24 (Figure 2.7). However, the expression difference of starch metabolic genes was not obviously correlated with different ploidy hybrids. In general, starch content was increased in the diploid, triploid and tetraploid hybrids, with the highest starch content in the tetraploid hybrid (Col4XC24-4) (Figure 2.5D). The starch content was significantly different in the reciprocal hybrids at diploid, triploid, and tetraploid levels, with one exception. Similar starch content was observed in Col4XC24-2 and C24-2XCol4 triploid hybrids, although the biomass and seed size were dramatically different between these crosses (Figures 2.3 and 2.4).

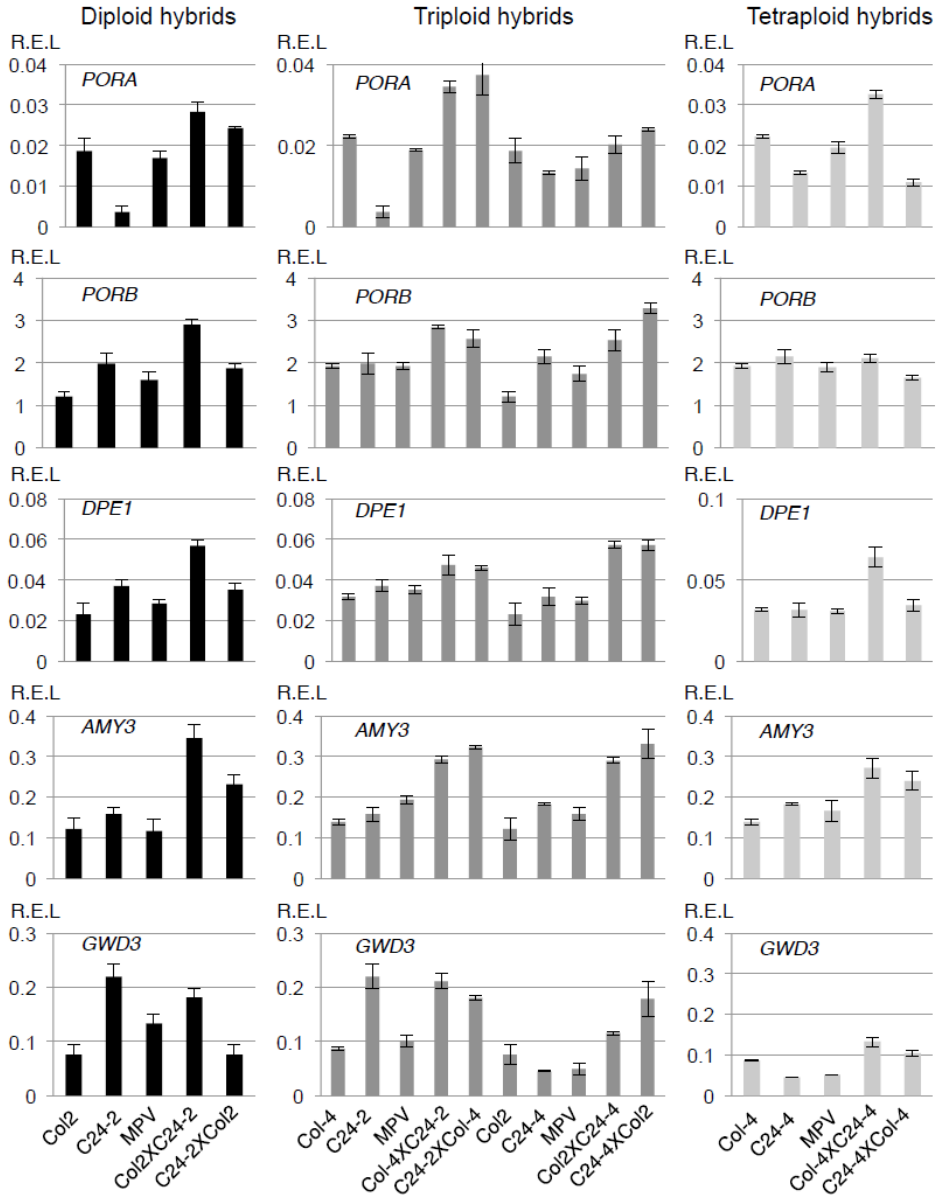


Figure 2.7. Expression of the genes involved in chlorophyll and starch metabolism in ColXC24 hybrids and parents at ZT6.

Discussion

Analysis of phenotypic variation and dosage regulation in some *Arabidopsis* intraspecific hybrids at diploid, triploid and tetraploid levels has shown several interesting findings. First, all hybrids tested, except for the triploid hybrids with a tetraploid mother, display biomass vigor compared to their respective parents. The highest biomass was found in the triploid hybrids with a diploid mother and a tetraploid father, and this maternal effect is dependent on genotypes (e.g., C24 vs. Col). Additionally, there appears to be a dosage compensation mechanism for ploidy effects on plant size, with tetraploids being the same size as diploids. In maize, the plant size increases from haploid to triploid, but decreases in the tetraploid [104], which is consistent with smaller *A. thaliana* plants in the haploids than in the diploids [116, 117].

Second, polyploids have bigger cells with relatively low density, which supports the previous notion of estimating polyploid plant frequency using stomatal size in fossil samples [118]. Stomatal size is positively correlated with ploidy levels, independent of hybridization effects except in the diploid hybrids, whereas stomatal density is negatively correlated with genome dosage, with a few exceptions probably because of genotypic effects. The data are consistent with that in *A. thaliana* and *A. lyrata* diploids and tetraploids [119]. No significant difference in stomatal cell size in the hybrids may suggest that hybrid plants increase cell number probably resulting from rapid cell expansion. However, this notion needs to be confirmed with a precise measure of cell size using other methods such as cell sorting.

Third, seed and flower size do not compensate for increased levels of ploidy, opposed to what is seen in vegetative tissues, because tetraploid seeds are bigger and heavier than diploid and triploid seeds, as observed in other plants species [120]. Seed size is also increased in the hybrids relative to the parents in the same ploidy level.

Modest seed size increases have also been observed in ColXC24 and CviXL*er* F1 hybrids [9, 30, 121] and in tomato hybrids [122]. However, the effects of hybridity on seed size and weight in tetraploids are not obvious, possibly because tetraploids already produce seeds at the maximum size, so no further increase due to hybridization is possible. It would be interesting to examine seed size variation in increased ploidy levels such as hexaploid and octoploid plants.

A typical seed contains a diploid embryo and a triploid endosperm as a result of double fertilization in angiosperms [123]. The proper seed development requires an endosperm ratio of 2m:1p (maternal:paternal) genomes [123, 124]. In *A. thaliana*, increasing the paternal genome ratio (2m:2p) in paternal-excess triploids between a diploid "mother" and a tetraploid "father" (2X4) produces larger seeds. In contrast, increasing the maternal genome ratio (4m:1p) in maternal-excess triploids between a tetraploid mother and a diploid father (4X2) lead to smaller seeds [124, 125]. This phenomenon is consistent with the parental conflict model that explains genomic imprinting in mammals [126, 127]. In contrast to imprinting in embryo, the endosperm is not genetically transmissible to offspring. In seeds, triploid endosperm is a maternal tissue that provides the nutrient reserve for seedling growth, as placenta do in mammals, except that the latter is a diploid tissue. It is notable that larger seeds in these triploid hybrids with a diploid mother and tetraploid father also lead to increased biomass during early stages of vegetative growth. The molecular basis for this link remains to be investigated. One of the triploid hybrids (C24XCol4) produces greenish and viviparous seeds, which is probably related to expression disruption in genes such as *TTG2* in the seed coat in a dosage and genotype-dependent manner [128].

Fourth, expression of circadian clock genes is altered by hybridization irrespective of the ploidy levels, and genome dosage has no obvious effects on the overall

level of clock gene expression. There is an overall correlation of *CCA1* repression and *TOC1* upregulation with increased levels of downstream genes, starch and chlorophyll in hybrids at all ploidy levels tested. The degree of clock gene expression changes is increased from intraspecific hybrids to interspecific hybrids or allotetraploids [63], which is consistent with higher biomass vigor and starch content in interspecific hybrids than in intraspecific hybrids. Compared to genome dosage, genetic distance and hybridization are critical in determining the level of biomass heterosis. However, expression of some transgenes [129] and endogenous genes [25, 32] is dependent on genome dosage. This suggests that ploidy effects on some genes may not be directly related to obvious growth and developmental phenotypes. Expression of clock genes is probably compensated by changes in genome dosage.

Finally, changes in clock gene expression and starch content are different in reciprocal crosses at the diploid and tetraploid levels. This parent-of-origin effect appears to depend on maternal parents. However, in the triploid hybrids, the effects on starch content in the reciprocal crosses are opposite to the effects on biomass and seed size. This suggests that the effects of genome dosage and hybridization on seeds and overall plant size are uncoupled from starch metabolic pathways. In the reciprocal hybrids, upregulation of clock-regulated downstream genes is not always obvious, and there is less upregulation of downstream genes in intraspecific F1 hybrids than in allotetraploids [63]. This indicates that there may be additional mechanisms or downstream effects in the F1 hybrids than in the allopolyploids. For example, small RNAs are shown to be differentially regulated in *Arabidopsis* allopolyploids [65] and *A. thaliana* hybrids [13] relative to the parents, and DNA methylation and gene expression changes are related in rice intraspecific hybrids [29]. Together, these data should offer new insights into a better understanding of the complexity of growth vigor in intraspecific and interspecific

hybrids, as well as in allopolyploids that are of direct relevance to plant evolution and crop production.

Chapter 3. A role for CHH methylation in the parent-of-origin effect on altered circadian rhythms and growth vigor in plant hybrids

Background and rationale

The growth vigor in *Arabidopsis* hybrids and allotetraploids is partly controlled by epigenetic regulation of central circadian clock oscillator components including *CCA1* [31, 63, 75], although the mechanism for regulating maternal and parental alleles of clock genes that influence heterosis remains unknown. In *A. thaliana*, 24 nt siRNAs generated by RNA polymerase IV, encoded by *NRPD1a*, guide de novo methylation of CHH and CHG sites (H = A, T, or C) through a process called RNA-directed DNA methylation (RdDM) [130-134]. In this process, siRNAs are generated by the endoribonuclease DCL3 which are then loaded onto ARGONAUTE 4 (AGO4), and AGO4-bound siRNAs guide the cytosine methyltransferase activity of DOMAINS REARRANGED METHYLASE 2 (DRM2) [133, 135, 136]. Maintenance of CG and CHG DNA methylation requires DNA METHYLTRANSFERASE I (MET1) and CHROMOMETHYLASE 3 (CMT3), respectively, and CHH methylation requires continual de novo methylation by DRM2 [133, 137]. DNA methylation is also known to affect circadian gene expression. *CCA1* and *LHY* were upregulated in *A. thaliana* DNA methylation mutants [138] and in plants treated with 5'-aza-2'deoxycytidine, a DNA methylation inhibitor [31].

Changes in *CCA1* expression and DNA methylation were observed in *A. thaliana* hybrids [31], but the relationship between them is unclear. As discussed in Chapter 2, the parent-of-origin effect on biomass vigor in reciprocal *A. thaliana* hybrids suggests an epigenetic cause [75]. In allotetraploids, the maternally transmitted *A. thaliana CCA1* allele is more repressed than the paternally transmitted *A. arenosa* allele, which is associated with histone modifications [63]. Other studies have also documented reciprocal size differences in hybrids derived from C24 and L*er* [13, 31]. The available

data suggest that changes in methylation levels at the *CCA1* locus are different in reciprocal crosses, or that background methylation levels are different in reciprocal crosses, thus indirectly influencing *CCA1* expression. Here the hypothesis that altered expression of circadian genes in hybrids is mediated by epigenetic factors, such as DNA methylation, was tested. In a collaborative effort between myself and other scientists (if any results are presented that were generated by a collaborator, this will be made clear in the text), we investigated how and when the parent-of-origin effect on *CCA1* expression and growth vigor is established in *A. thaliana* hybrids. The results suggest that AGO4-mediated changes in CHH methylation affect *CCA1* expression and growth vigor in hybrids during vegetative and embryo development.

Materials and methods

Plant growth and materials

Plant materials included three *Arabidopsis thaliana* ecotypes, C24, Landsberg *erecta* (L*er*), and Columbia (Col-0), three mutants in DNA methylation and small RNA biogenesis genes, including *met1-1* (L*er*) [139], *ddm1-2* (L*er*) [140] and *ago4-1* (CS6364, L*er*). *met1-1* and *ddm1-2* mutant seeds were kindly provided by Eric Richards at the Boyce Thompson Institute for Plant Research, and *ago4-1* was obtained from the *Arabidopsis* Biological Resource Center (ABRC). For comparison in F1 hybrids and crosses, manual pollination was used to produce seeds in both parents and reciprocal hybrids. For gene expression and starch analyses in vegetative tissues, plants were grown for 3 weeks in 16/8 hr (light/dark) cycles at 22/18ºC and harvested at given zeitgeber time (ZT0 = dawn). Rosette leaves were harvested from a pool of 6-12 plants as one biological replicate and used immediately or frozen in liquid nitrogen for future use. Leaves were

collected prior to bolting to minimize developmental variation among genotypes. Except where noted otherwise, three replicates were used for each experiment.

For gene expression analysis in developing siliques, manual pollination was performed one day after emasculation. Young siliques at 5 days after pollination (DAP) were harvested every 3 hours (h) for a period of 24 h for diurnal expression analysis.

Transgenic plants expressing luciferase reporter

The *CCA1::LUC* construct was transformed into L*er* as previously described [141] to generate L*er*(*CCA1::LUC*) or L*er*C stable transgenic plants for this study. To generate the C24(*CCA1::LUC*) or C24C line, a *CCA1* promoter (from -715to -1 bp, relative to the transcription start site plus full 5' UTR) was amplified by PCR and cloned into the plasmid between the restriction sites *Xho*I and *Nco*I. A *CCA1::LUC* plasmid construct was generated by inserting the luciferase gene between the restriction enzyme sites *Nco*I and *Bam*HI in the pFAMIR plasmid that was modified from pFGC5941 [142]. The construct was introduced into *A. thaliana* (C24) plants using *Agrobacterium* mediated transformation (*A. tumefaciens* GV3101) with the floral dip method [143]. Primary transformants (seedlings) were screened on Murashige and Skoog (MS) agar medium (M9274, Sigma-Aldrich, St. Louis, MO) supplemented with 7.5 μg mL−1 of Basta (Sigma, St. Louis, MO). Stable transgenic plants (T2 and later) with a uniform herbicide resistance were used for the expression assays and for making crosses.

Embryo dissection and culture (developed by Dr. Helen H. Yu)

Siliques 10 days after pollination (DAP) were harvested and rinsed with 70% ethanol and soaked in 100% Chlorox for 2 minutes. After rinsing with autoclaved water twice, the siliques were kept in sterilized liquid Murashige and Skoog medium in a Petri dish. Embryos were dissected under an optical microscope (SMZ445, Nikon, Melville,

NY) and transferred to a plate containing agar embryo culture medium, which contained 40% sucrose, 0.5X Murashige and Skoog salts, 0.9 mg/L thiamine, 0.5 g/L 4-morpholineethanesulfonic acid sodium salt (MES) (Sigma, St. Louis, MO), 8 g/L agar, and 0.69g Leu-DO amino acid supplements (Clontech, Mountain View, CA). Final pH was adjusted to 5.9 with KOH. Forty to fifty embryos from each genotype were transferred to one agar plate and cultured in an incubator at 22ºC (16/8-hr light/dark cycles) for 2 days. A total of 24 healthy embryos (no brown spots or any visible damage) from each genotype were transferred into a 96-well microtiter plate (Nagle Nunc International, Rochester, NY), containing 40% sucrose, 0.5X Murashige and Skoog salts, 0.9mg/L thiamine, 0.5g/L MES (Sigma, St. Louis, MO), 8g/L agar and 0.69g Leu-DO amino acids supplements. After adding luciferin to a final concentration of 2.5 mM, the plate was subjected to luciferase assays for a period of 5-7 days (see below).

Luciferase assays and data analysis

Stable transgenic embryos or seedlings containing *CCA1::LUC* constructs were analyzed using a TopCount NXT luminometer and scintillation counter (PerkinElmer, Shelton, CT). For seedlings, seeds were sterilized and plated on 1% (w/v) agar MS media plus 30 g/L of sucrose. Seeds were stratified 2 days (d) in the dark at 4°C and then transferred into 16-h light/8-h dark (LD) cycle for 8 d at 22°C. Seedlings were transferred to opaque microtiter plates (Nunc, Denmark) containing agar MS medium plus 30 g/L sucrose, and then 30 μL of 0.5 mM luciferin (Gold Biotechnology, St. Louis, MO) was added to each well. Microtiter plates were covered with clear plastic MicroAmp sealing film (AppliedBiosystems, Foster City, CA), in which holes were placed above each well for seedling gas exchange. Plates were moved to the TopCount and interleaved with three clear plates to allow light diffusion to the seedlings. Luciferase activity was measured

approximately every 1 h by integrating photons emitted by seedlings during a 10-second sampling period. Data was analyzed by importing data into the Biological Rhythms Analysis Software System (BRASS) excel macros (available from http:www.amillar.org).

All values are presented as mean ± standard error of the mean (SEM). Expression peaks were shown as bioluminescence counts. Unless noted otherwise, each data point was averaged from between 24 to 32 plants in each experiment, and graphic data from one of three replicated experiments were shown.

RNA preparation and reverse transcription-polymerase chain reaction (RT-PCR)

Total RNA was extracted using Plant RNA reagent (Invitrogen, Carlsbad, CA). The first strand cDNA synthesis was performed using RT Superscript III (Invitrogen, Carlsbad, CA). An aliquot (1/100) of cDNA was used for quantitative RT-PCR (qRT-PCR) analysis using the primer pairs (Table 3.1) in an ABI7500 machine (Applied Biosystems, FosterCity, CA). Amplification of *ACT7* (At5g09810) served as a control to estimate relative expression levels (R.E.L.).

Table 3.1 Primers for qRT-PCR and methylation assays

qRT-PCR primers	Sequence (5'-3')
CCA-F	CCGCAACTTTCGCCTCAT
CCA1-R	GCCAGATTCGGAGGTGAGTTC
LHY-F	GGGACAAAGACTGCTGTTCAGAT
LHY-R	TTTGTGAAGAACTTTTGTGCATGA
TOC1-F	GTTGATGGATCGGGTTTCTC
TOC1-R	TCATGACCCCATGCATACAG
CHE-F	TAATGGGTGGTGGTGGTTCTG
CHE-R	GCAAAGCTCCAGACTTGTCC
ACT7-F	AGATTCTCACTGAGAGAGGTTACATGTTCA
ACT7-R	ATGACTTGTCCATCAGGTAGCTCG
Bisulfite sequencing primers	
BS1-At-382CCA1p F	TTAGAGTGTGAGAATAGYGYGTGTA
BS2-At-39CCA1p R	RCTACAAAARAAACTRCCATA
A1B-F	GGAATTCACCAACCAAATCTCCTTCC
A1B-R	GCTCTAGATAGYAAGAAYAATAGGAAGAG

Genomic DNA extraction and bisulfite sequencing

Genomic DNA was extracted from 3 week old seedlings (100 mg) using the DNeasy Plant mini Kit (Qiagen, Valencia, CA). About 500 - 800 ng genomic DNA was then used for bisulfite conversion using the EpiTect bisulfite Kit (Qiagen) according to the manufacturers' instructions. Bisulfite treated DNA (5 µL) was then amplified by PCR in a 25 µL reaction using ZymoTaq DNA polymerase (ZYMO Research Corporation,

Irvine, CA) and degenerate primers (Table 3.1) targeting the -382/-39 *CCA1* promoter region containing the G-box and CHE-binding site (a motif region spanning -280/-230), or a region of *ASA1* which was used as a bisulfite conversion control [140]. PCR products were then resolved in a 1% agarose gel, excised and purified using UltraClean DNA purification Kit (Mo Bio Laboratories, Carlsbad,CA), and cloned into pGEM-T vector (Promega Corporation, Madison, WI) for sequencing. For each plant genotype, at least 16 independent top-strand clones were sequenced. Bisulfite DNA sequences and the levels of DNA methylation at the *CCA1* promoter were analyzed using the online program Kismeth [144]. For each genotype, the percentage of cytosine methylation at each context (CG, CHH or CHG) was calculated and differences in DNA methylation between two genotypes were analyzed using the student t-test.

Biomass analysis

Whole rosettes from hybrids and parents were harvested at approximately 3 weeks of age (before bolting) and placed in Lawson #217 hybridization bags (Lawson Bags, Northfield, Illinois). The weight from aerial rosette leaves was determined after drying the plants at 80°C for 24 h. Aerial rosettes of 6 plants in three biological replicates were weighed individually, and the average was used to calculate standard deviations [75].

Results

Parent-of-origin effects on biomass heterosis in hybrids

Reciprocal F1 hybrids between *A. thaliana* C24 and L*er* displayed biomass heterosis [9, 13, 31, 75]. The biomass (dry weight) was also significantly higher in C24XL*er*C (L*er*C is L*er* containing *CCA1:LUC* transgene) than in L*er*CXC24 hybrids (Figure 3.1B,C; $p < 0.05$), but this was not observed in the control crosses L*er*CXL*er* and

L*er*XL*er*C (Figure 3.1A,D). This parent-of-origin effect on biomass vigor was also observed in other reciprocal F1 hybrids between Col and C24 or between Col and L*er* [75], although the vigor level was lower in the latter. The variable degree of biomass vigor among different hybrids may reflect genotypic effects [4, 9]. The biomass difference between reciprocal hybrids between C24 and L*er* is consistent with starch accumulation being high in C24XL*er* than in L*er*XC24 (data not shown, work done by Dr. Eun-Deok Kim). Collectively, these data suggest a parent-of-origin effect on biomass heterosis. To minimize potential genotypic effects, unless noted otherwise, further analyses were performed in F1 crosses between C24 and L*er* ecotypes including several mutants in the L*er* background.

Parent-of-origin effects on circadian rhythms

Altered *CCA1* expression correlated with growth vigor in allopolyploids, hybrids, and diploids [63]. Repressing *CCA1* peaks in *TOC1:cca1*(*RNAi*) transgenic plants during the day increases starch content and biomass, while overexpressing *TOC1:CCA1* in transgenic plants decreases starch content and biomass. These data suggest an important role for altered *CCA1* expression in promoting growth vigor. Using qRT-PCR to analyze *CCA1* expression in reciprocal hybrids, it was found that lower *CCA1* expression levels were correlated with higher levels of starch and biomass in C24XL*er* than in L*er*XC24 hybrids (data not shown, analysis performed by Dr. Danny Ng and Dr. Grace Kim), suggesting an anti-correlation between endogenous *CCA1* expression levels and growth in hybrids.

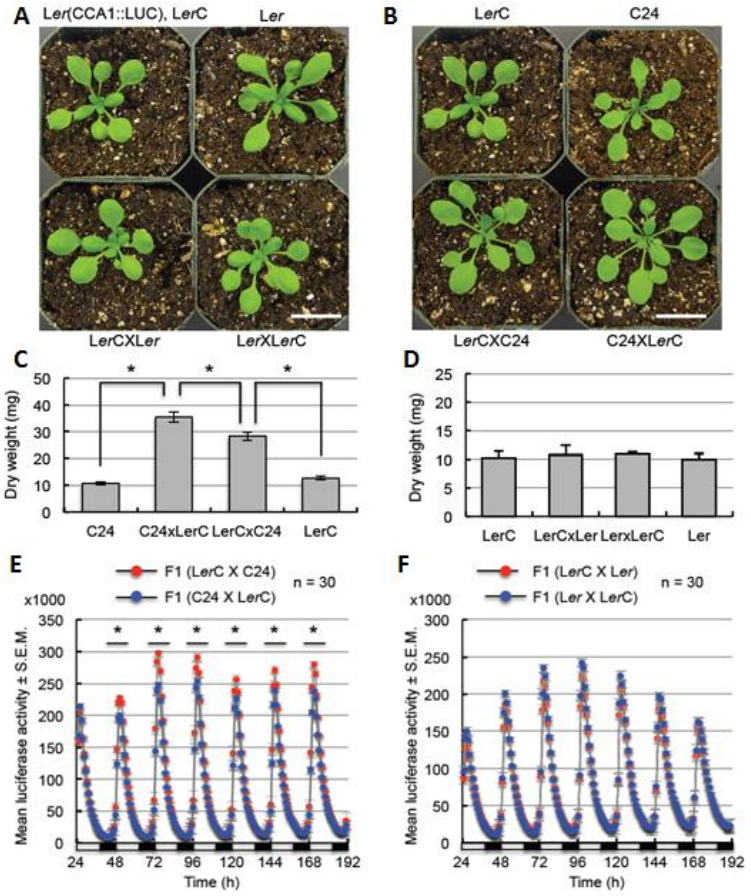


Figure 3.1 Parent-of-origin effects on *CCA1* expression and biomass accumulation in hybrids

(**A, B**) Photos of seedling plants in L*er*(pro*CCA1*:LUC) or L*er*C, L*er*, L*er*CXL*er*, and L*er*XL*er*C (**A**) and C24, L*er*C, L*er*CXC24, and C24XL*er*C (**B**). Scale bar = 20 mm. (**C**) Higher aerial dry weight (Y-axis) of 3-week old seedlings in F1(C24XL*er*C) than in F1(L*er*CXC24), relative to the parents C24 and L*er*(CCA1:LUC) or L*er*C ($p < 0.05$). (**D**) No difference in aerial dry weight (Y-axis) in 3-week old seedlings of reciprocal crosses between L*er* and L*er*C. (**E**) Mean values (± standard errors, s.e.m.) of bioluminescence counts (in thousands, Y-axis) in the seedlings of the reciprocal hybrids L*er*CXC24 (red) and C24XL*er*C (blue). Each data point was averaged from 30 plants with standard errors. Black lines with asterisks indicate the range of time points with statistically significant differences between the reciprocal crosses ($p < 0.05$). X-axis: hours with an alternating cycle of light (open box) and dark (dark box). (**F**) Bioluminescence (Y-axis, in thousands, ± s.e.m.) of seedlings in the reciprocal F1 crosses L*er*CXL*er* (red) and L*er*XL*er*C (blue). Same notes as in (**E**).

To further examine differences between reciprocal hybrids, the L*er*C transgenic line was employed [141]. Using the reporter line, diurnal oscillation of *CCA1* expression in a period of 5-7 days was examined in hybrids. Unless noted otherwise, bioluminescence assays included three biological replicates each with 24-32 seedlings,

and resulting data points at approximately 1-hour intervals in 5-7 days were analyzed for statistical significance using paired t-tests between each comparison (e.g., reciprocal hybrids). In reciprocal F1 hybrids between C24 and L*er*C, the *CCA1* expression peak was statistically significantly higher when *CCA1:LUC* was transmitted through the maternal rather than the paternal parent (Figure 3.1E, $p < 0.05$). Higher *CCA1* expression peaks in L*er*CXC24 than in C24XL*er*C inversely correlated with starch and biomass levels in the hybrids (Figure 3.1C). In the control crosses between L*er* and L*er*C lines, *CCA1:LUC* was equally expressed through the paternal or the maternal parent (Figure 3.1F), which had similar biomass (Figure 3.1D). These data indicate parent-of-origin effects on expression of the transgene *CCA1:LUC* and likely endogenous *CCA1*, which negatively correlates with biomass heterosis.

CHH methylation and AGO4 affect parent-of-origin effects on circadian gene expression

Parent-of-origin effects are often associated with paternal and maternal inheritance of DNA methylation patterns [145-147], and methylation levels in promoter regions correlate with gene expression levels [148]. To test this, methylation levels of CG, CHG, and CHH sites in the *CCA1* promoter region were examined using the bisulfite sequencing method [144]. Degenerate primers flanking the *CCA1* promoter region (-382/-39, relative to the transcription start site of +1) were used to amplify bisulfite-treated DNA, which was subsequently cloned and sequenced. Methylation levels of CHH, CHG, and CG sites were calculated and compared between reciprocal hybrids. In addition to this larger promoter fragment (-382 to -39), methylation analyses were also performed on a smaller region including a motif domain (-280 to -230) that contains a G-box and a *CCA1 HIKING EXPEDITION* (*CHE*), a Class I TCP protein, -binding site (TBS) [149], which is named GTBS (Figure 3.2A). Only quantified results from the GTBS region are

displayed here. This GTBS region was selected for the methylation analysis because *CHE* binds there to mediate *CCA1* expression [149]. To ensure that bisulfite conversion efficiency between reciprocal hybrids was equal, an unmethylated region of the *ASA1* locus was cloned and sequenced (Figure 3.1A) [140]. Methylation levels at *ASA1* were low in all genotypes and were not significantly different between reciprocal crosses, indicating near-complete conversion of the DNA (Figure 3.1B-D). In reciprocal F1 hybrids between C24 and L*er* (wild-type), methylation levels of CG, CHG and CHH sites were generally higher in C24XL*er* hybrids than in L*er*XC24 hybrids in the larger *CCA1* promoter region, although such differences were statistically insignificant. However, within the GTBS region, CHH methylation levels were statistically significantly higher in C24XL*er* hybrids than in L*er*XC24 hybrids (Figure 3.2B,C) ($p < 0.05$), whereas CG or CHG methylation levels showed large variation and were in the opposite trend. This variability could result from few CG and CHG sites, compared to CHH sites in the promoter regions analyzed (Figure 3.2B). The differences in CHH methylation in the GTBS region between the reciprocal F1 hybrids of C24 and L*er* were correlated negatively with the endogenous *CCA1* expression and positively with biomass and starch content in F1 hybrids between C24 and L*er*.

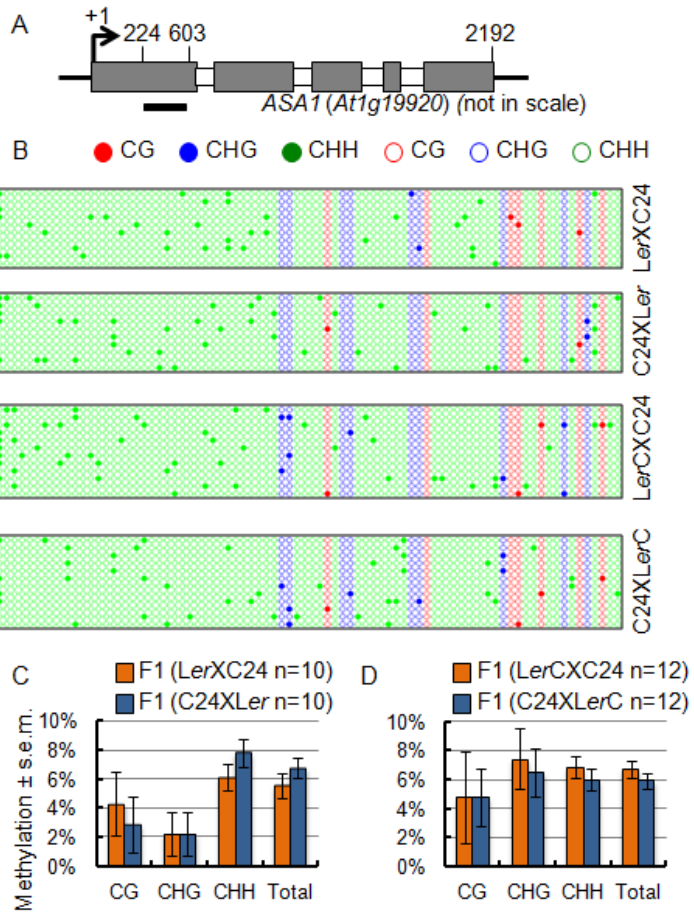


Figure 3.1 Bisulfite sequencing analysis of DNA methylation at *ASA1*

(**A**) Schematic diagram of the *ASA1* locus (*At1g19920*) and the targeted region (224-603; black bar) used as the positive control of bisulfite conversion. (**B**) Dot-plot analysis of methylation in the indicated hybrids. Red, blue, and green circles indicate CG, CHG, and CHH methylation (filled) or no methylation (open). (**C**) Percentage of methylation changes (± s.e.m.) of the *ASA1* locus between reciprocal hybrids of L*er*XC24 (orange) and C24XL*er* (pale blue). (**D**) Percentage of methylation changes (± s.e.m.) of the *ASA1* locus between reciprocal hybrids of L*er*CXC24 (orange) and C24XL*er*C (pale blue). (n = number of clones sequenced).

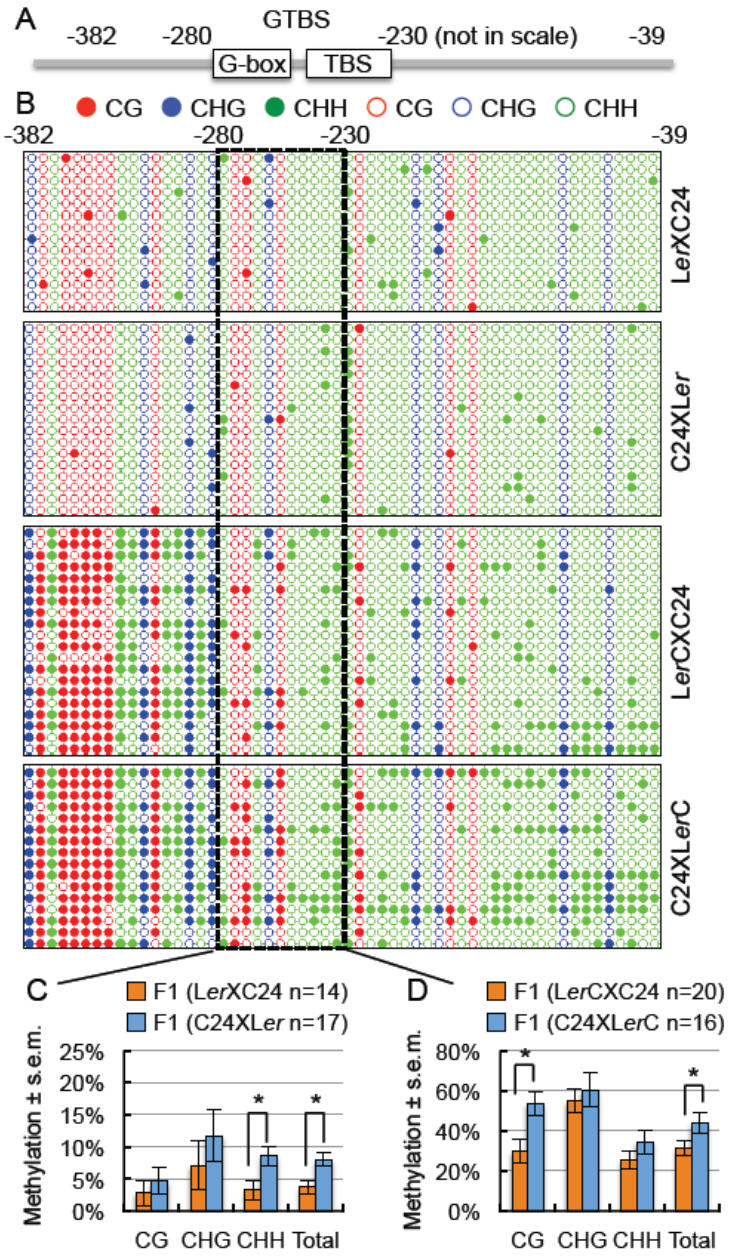


Figure 3.2 Bisulfite sequencing analysis of DNA methylation in reciprocal hybrids

(**A**) Diagram of a large promoter region (-382 to -39) of *CCA1* and a motif region (-280 to -230) that contains a G-box and a *CHE* with a Class I TCP protein-binding site (TBS), which is named GTBS. (**B**) Dot-plot analysis of CG, CHG, and CHH methylation changes in L*er*XC24 (top), C24XL*er* (second), L*er*CXC24 (third), and C24XL*er*C (bottom) in the large promoter region and GTBS region (boxed). A total of 14-20 individual promoter fragments were sequenced and analyzed in each sample. Red, blue, and green circles indicate CG, CHG, and CHH methylation (filled) or no methylation (open). (**C**) Percentage of methylation changes (± s.e.m.) in the GTBS region between reciprocal hybrids of L*er*XC24 (orange) and C24XL*er* (pale blue). (n = number of clones sequenced). (**D**) Percentage of methylation changes (± s.e.m.) in the GTBS region between reciprocal hybrids of L*er*CXC24 (orange) and C24XL*er*C (pale blue). (n = number of clones sequenced).

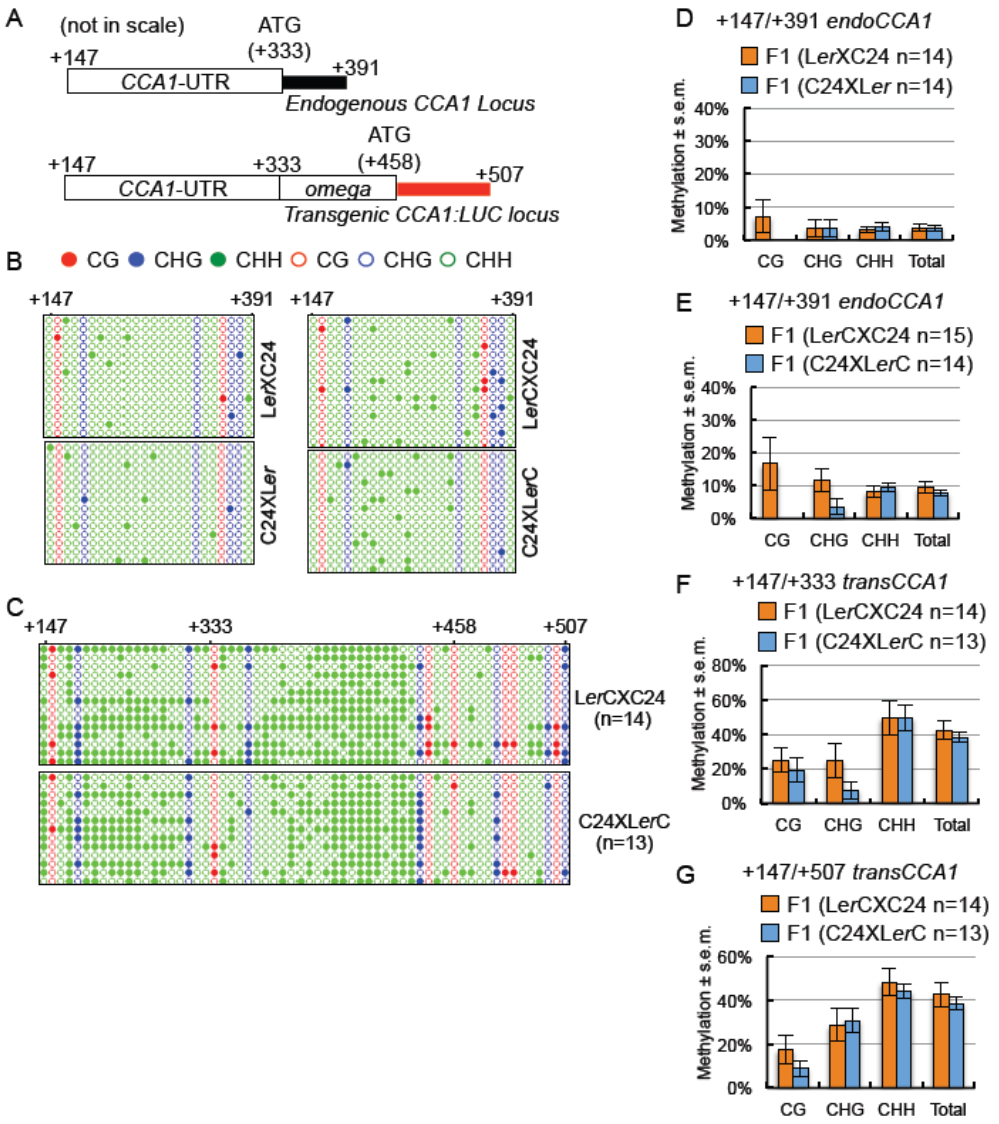


Figure 3.3 Bisulfite sequencing analysis of DNA methylation at the 5'UTR of *CCA1*

(**A**) Schematic diagram of the endogenous *CCA1* locus and transgene *CCA1:LUC* locus targeted for bisulfite sequencing PCR. The 5' leader of the tobacco mosaic virus (omega) in the transgenic *CCA1:LUC* was indicated. (**B-C**) Dot-plot analysis of CG, CHG, and CHH methylation in the indicated hybrids. Red, blue, and green circles indicate CG, CHG, and CHH methylation (filled) or no methylation (open). (**D**) Percentage of methylation changes (± s.e.m.) of the endogenous *CCA1* UTR region between reciprocal hybrids of L*er*XC24 (orange) and C24XL*er* (pale blue). (**E-F**) Percentage of methylation levels (± s.e.m.) of the endogenous *CCA1* UTR region (**E**) and the same region in the transgene *CCA1:LUC* locus (**F**) between reciprocal hybrids of L*er*CXC24 (orange) and C24XL*er*C (pale blue). (**G**) Percentage of methylation levels (± s.e.m.) of the *CCA1*-UTR and the 5' leader (omega) of tobacco mosaic virus in the transgene *CCA1:LUC* locus in L*er*CXC24 reciprocal hybrids. n = number of clones sequenced.

DNA methylation was further analyzed in reciprocal F1 plants between L*er*C (transgenic line) and C24 or L*er*C (Figure 3.2B,D). The overall methylation levels at the *CCA1* promoter region was higher in reciprocal hybrids containing both the endogenous *CCA1* and transgene *CCA1:LUC* than in the reciprocal crosses between wild-type plants (containing only the endogenous *CCA1* locus) (Figure 3.2C,D). The increased level of DNA methylation was probably related to the transgene *CCA1* promoter. In the hybrids between C24 and L*er*C, methylation levels at all sites in the larger *CCA1* promoter region were slightly higher in C24XL*er*C hybrids than in L*er*CXC24 hybrids. Within the GTBS, although CHH methylation levels in the endogenous *CCA1* promoter were statistically different between L*er* and C24 reciprocal F1 hybrids (Figure 3.2C), the trend of higher CHH methylation in C24XL*er*C hybrids than in L*er*CXC24 hybrids (involving the transgene) remained, but this difference was statistically insignificant, whereas CG methylation levels were significantly different (Figure 3.2D). However, total and CG methylation levels were statistically significantly higher in the C24XL*er*C than in the reciprocal cross (Figure 2D). This is probably because of an increased methylation level in the CG sites of the transgene promoter, which may obscure the difference in the CHH methylation (Figure 3.3). In control crosses between L*er*C and L*er*, methylation levels of all sites were similar in the larger *CCA1* promoter region or within the GTBS (not shown).

Higher methylation levels in the transgenic *CCA1:LUC* locus than in the endogenous *CCA1* locus were also observed in 5' untranslated region (UTR) (Figure 3.3A), in which endogenous and transgene loci could be discriminated in the F1 hybrids (Figures 3.3B-C). However, the methylation level differences between the reciprocal hybrids were not significant (Figure 3.3D-F). These data suggest that the transgene is highly methylated especially at CG and CHG sites, and the methylation changes in the 5'

UTR are not correlated with the transgene or endogenous *CCA1* expression in the reciprocal hybrids. Instead, CHH methylation changes in the GTBS region correlated with lower *CCA1* expression when C24 is the maternal parent in the hybrids.

In the RdDM pathway [133], AGO4 controls locus-specific methylation of CHH and CHG sites [135, 136]. *CCA1:LUC* expression was examined in the *ago4-1* mutant to determine if AGO4 regulates *CCA1*. Indeed, *CCA1:LUC* expression levels were statistically significantly higher in the *ago4-1* homozygous mutant (red) than in the wild-type (blue) (Figure 3.4A, $p < 0.05$). *CCA1:LUC* expression was then examined in reciprocal F1 crosses involving the *ago4-1* mutant. Remarkably, in the F1 crosses between the *ago4-1* (L*er*) mutant and L*er*C, *CCA1* expression peaks were statistically significantly higher when the *CCA1:LUC* allele was transmitted through the paternal (blue) than through the maternal (red) parent (Figure 3.4B). In the control crosses of L*er*CXL*er* and L*er*XL*er*C, *CCA1* expression peaks of maternal and paternal origins were equal (Figure 3.1F). Higher *CCA1* expression levels in the *ago4-1*XL*er*C cross than in the L*er*CX*ago4-1* cross correlated with reduced methylation levels of all sites ($p < 0.05$) in the former than in the latter in both the larger promoter region and the GTBS region (Figure 3.4D, methylation analysis performed by Dr. Danny Ng). In contrast, biomass was higher in the L*er*CX*ago4-1* cross than in the *ago4-1*XL*er*C cross, and the difference was narrowly statistically insignificant ($P = 0.052$) (Figures 3.4E and 3.5A). This suggests that other factors induced by the paternal or maternal *ago4* allele may also affect growth vigor. In reciprocal F1 hybrids between C24 and *ago4-1* (L*er*), the biomass in the F1 (*ago4-1*XC24) was increased, making it similar to but not higher than that in F1 (C24X*ago4-1*) (Figure 3.5C, E).

To test if maintenance of DNA methylation affects *CCA1* expression, reciprocal F1 crosses were made between L*er*C and *met1-1* (L*er*) [139] or *ddm1-2* (L*er*) [140]. In

F1 crosses between L*er*C and the *met1-1* mutant no alteration of the parent-of-origin effect on *CCA1* expression was found (Figure 3.4C), which had similar biomass (Figure 3.4F). A similar trend was seen in the reciprocal crosses between L*er*C and the *ddm1-2* mutant (*ddm1-2* results not shown). In reciprocal F1 crosses between C24 and *ddm1-2*, the parent-of-origin effect on biomass remained unchanged, namely, higher in C24X*ddm1-2* than in *ddm1-2*XC24 (Figure 3.5D, F). These data suggest that DDM1 and MET1, which affect mainly CG and CHG methylation but do not alter CHH methylation, do not change parent-of-origin effects on *CCA1* expression and biomass.

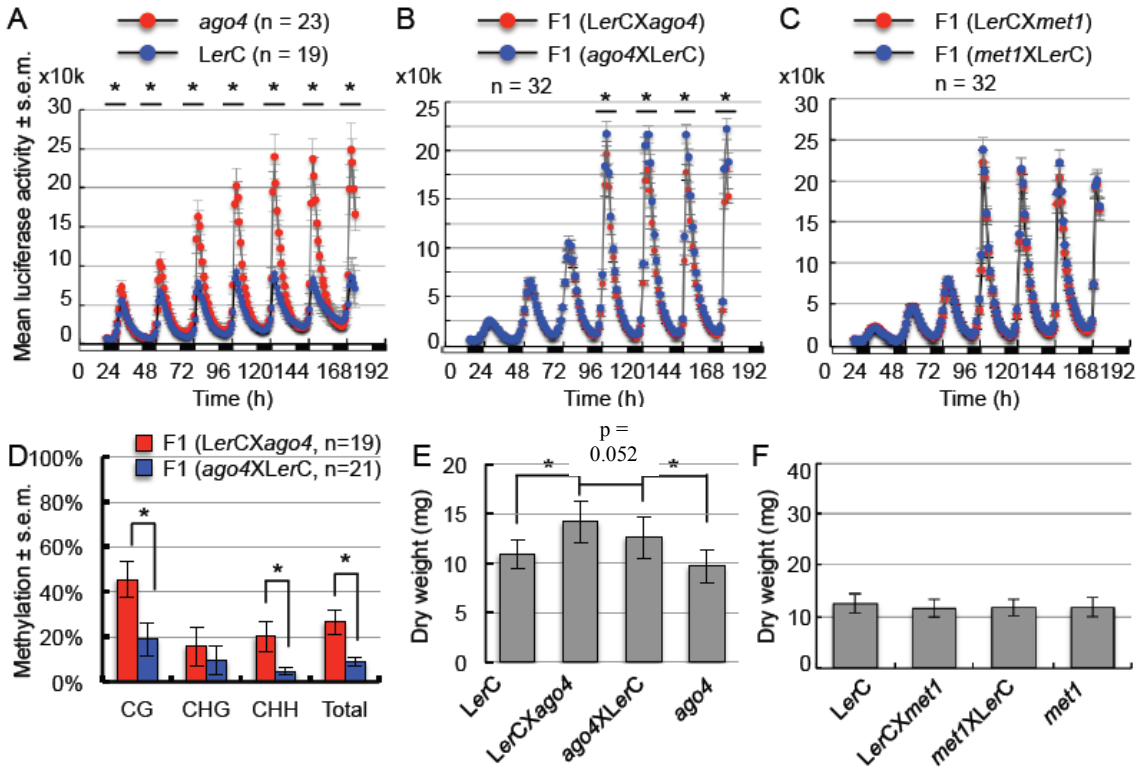


Figure 3.4 Parent-of-origin effects on *CCA1* expression depend on CHH methylation and AGO4 in reciprocal hybrids

(**A**) Mean values (± s.e.m.) of bioluminescence counts (in ten thousands, 10k, Y-axis) for *CCA1:LUC* expression in seedlings of the wild-type (L*er*C) (blue) and the *ago4-1* homozygous mutant (*ago4-1*C)(red). Black lines with asterisks indicate the peak time points with statistically significant differences between the reciprocal F1 crosses ($p < 0.05$). X-axis: hours with an alternating cycle of light (open box) and dark (dark box). (**B**) Mean values (± s.e.m.) of bioluminescence counts in the seedlings of the reciprocal F1 crosses between L*er*CX*ago4-1* (red) and *ago4-1*XL*er*C (blue). Same notes as in (**A**). (**C**) Mean values (s.e.m.) of bioluminescence counts in the seedlings of the reciprocal F1 crosses between L*er*CX*met1-1* (red) and *met1-1*XL*er*C (blue). (**D**) Percentage of methylation changes in the GTBS region between reciprocal F1 crosses of L*er*CX*ago4-1* (red) and *ago4-1*XL*er*C (blue). Asterisks indicate statistical significance ($p < 0.05$, n = number of clones in each replicate). (**E**) Slightly higher aerial dry weight (Y-axis) of 3-week old seedlings in L*er*CX*ago4-1* than in *ago4-1*XL*er*C crosses ($p = 0.052$) and relative to the parents *ago4-1* and L*er*. (**F**) No difference in aerial dry weight (Y-axis) in 3-week old seedlings of reciprocal F1 crosses between *met1-1* and L*er*C ($p = 0.7$).

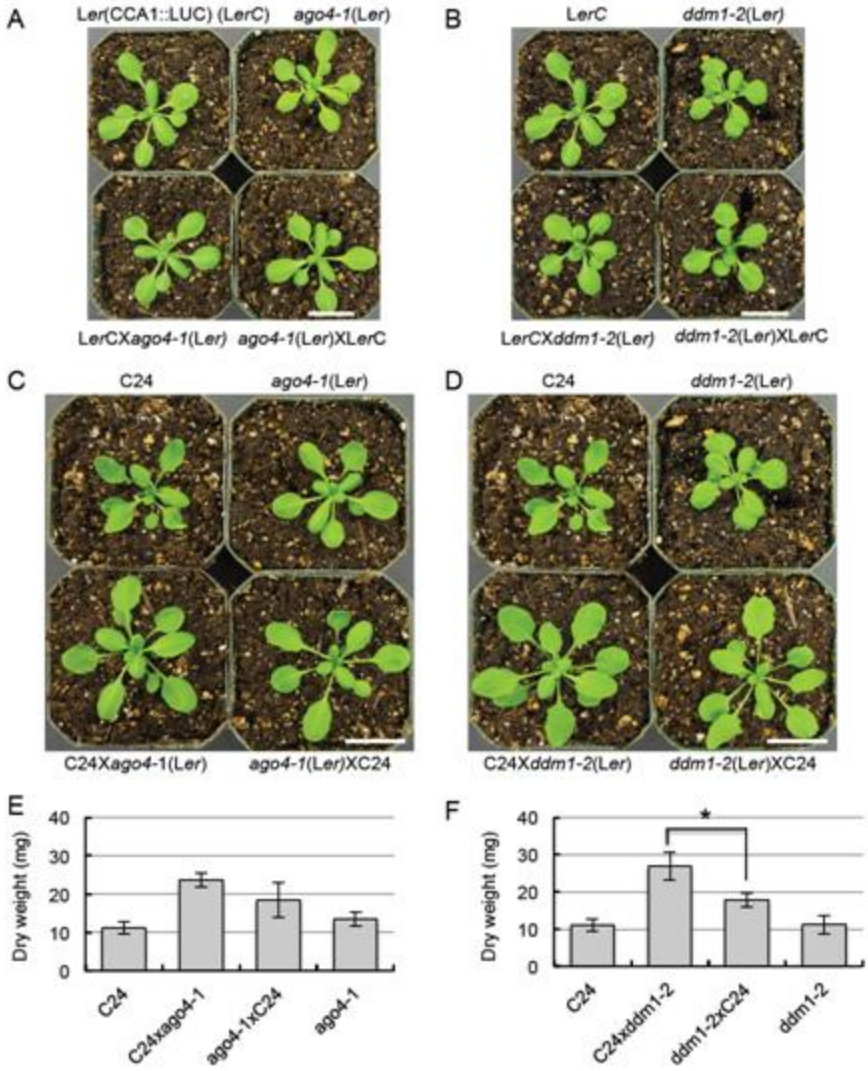


Figure 3.5 Biomass analysis in reciprocal hybrids and their parents

(**A**) Photos of typical seedling plants in L*er*C, *ago4-1*(L*er*), L*er*C X *ago4-1*, and *ago4-1* X L*er*C. (**B**) Photos of seedling plants in L*er*C, *ddm1-2*(L*er*), L*er*C X *ddm1-2*, and *ddm1-2* X L*er*C. Scale bar = 20 mm. (**C-D**) Photos of seedling plants in C24, *ago4-1*(L*er*), C24X*ago4-1*, and *ago4-1*XC24 (**C**) and C24, *ddm1-2*(L*er*), C24X*ddm1-2*, and *ddm1-2*XC24 (**D**). Scale bar = 20 mm. (**E-F**) Dry weight (mg) of above ground plants in C24, *ago4-1*(L*er*), C24X*ago4-1*, and *ago4-1*XC24 (**E**) and in C24, *ddm1-2*(L*er*), C24X*ddm1-2*, and *ddm1-2*XC24 (**F**).

Parent-of-origin effects on *CCA1* expression during early stages of embryo development

A key question is: when is the parent-of-origin effect on circadian rhythms established? The circadian clock is obviously present in leaves, but also in roots and shoots [150], and germinating seeds [151]. This suggests that the clock may function during embryo and seed development and that preferential expression of the maternal *CCA1* allele may be established during these stages. We first examined expression of *CCA1*, *LHY*, *TOC1*, and *CHE* in developing siliques 5 days after pollination (DAP) in two ecotypes (L*er* and Col-0). The mRNA abundance of the two morning-phased genes, *CCA1* and *LHY*, peaked at ZT0 and rapidly decreased towards a minimum at ZT12 and then rapidly increased towards ZT24 (Figure 3.6). The evening-phased genes, *TOC1* and *CHE*, exhibited antiphasic diurnal expression patterns compared to that of *CCA1* and *LHY*. The above data suggest that a robust clock is maintained in developing siliques as in leaves, roots [150], and germinating seeds [151].

In addition to an embryo, a typical seed contains an endosperm and a seed coat, which are maternal tissues that do not transmit genetic information to the next generation. To test if the clock is functional in developing embryos, we dissected embryos 10 days after pollination (DAP) (method developed by Dr. Helen H Yu, and all embryo experiments performed in collaboration with Dr. Yu), when the embryos could grow in culturing media (Figures 3.7A,B and 3.8A,B). After 4 days on culture media, embryos were subjected to bioluminescence assays (Figures 3.7C and 3.8C-E). Consistent with circadian rhythms in seedling leaves, *CCA1:LUC* expression peaks were statistically significantly higher when it was transmitted through the maternal parent (in L*er*CXC24) than through the paternal parent (in C24XL*er*C) (Figure 3.7C), which correlated negatively with embryo size in the former than in the latter (Figure 3.7A,B). In control

crosses between L*er*C and L*er*, the embryos had similar size (Figure 3.8A), and no *CCA1:LUC* expression difference was observed in the embryos of reciprocal F1 hybrids (Figure 3.8C). The stronger maternal expression of *CCA1* was also found in the embryos of another pair of reciprocal F1 hybrids between L*er* and C24(*CCA1:LUC*) or C24C transgenic line (Figure 3.8B,E), while in embryos of the control crosses, *CCA1:LUC* expression levels were equal (Figure 3.8D).

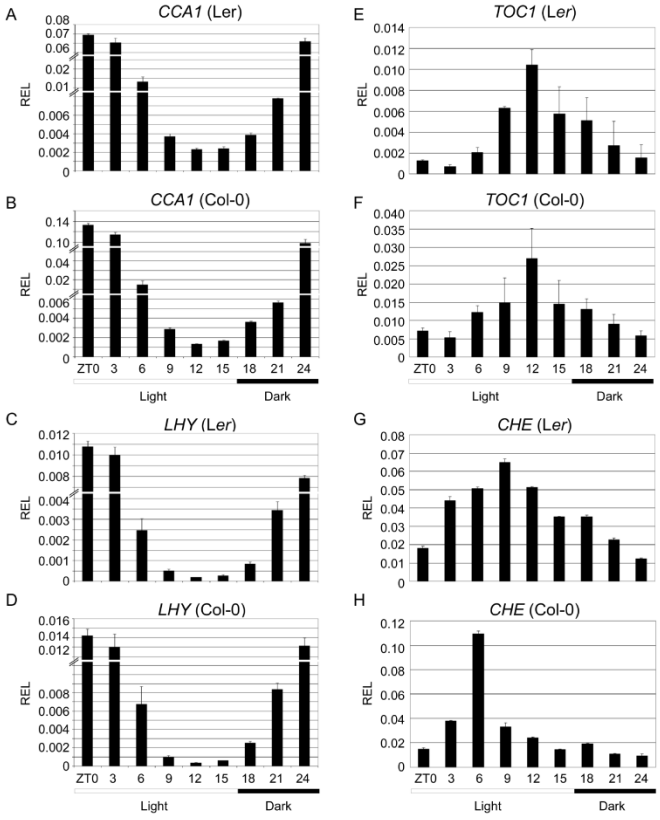


Figure 3.6 Diurnal expression of clock regulators in developing siliques in *A. thaliana* (Col-0 and L*er*)

Diurnal expression of *CCA1* in L*er* (**A**) and Col (**B**). Expression of *LHY* in L*er* (**C**) and Col (**D**). Expression of *TOC1* in L*er* (**E**) and Col (**F**). Expression of *CHE* in L*er* (**G**) and Col (**H**). R.E.L.: relative expression level of qRT-PCR results from three replications.

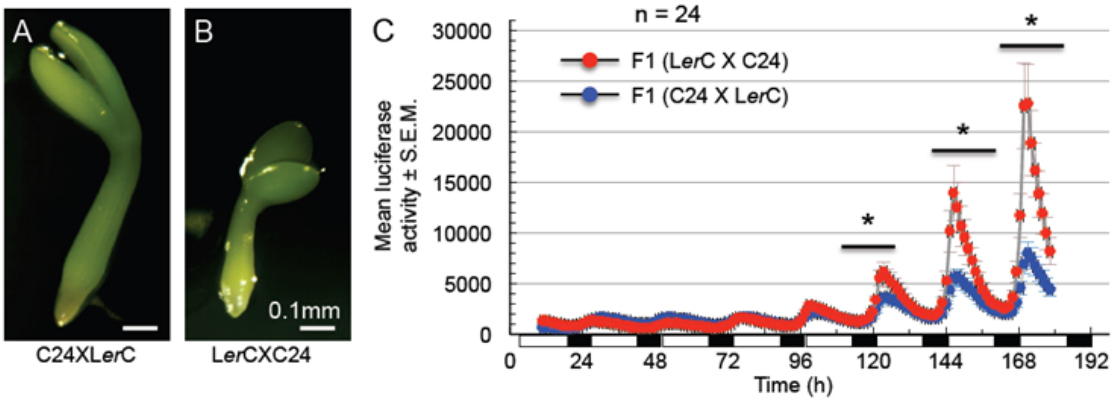


Figure 3.7 Parent-of-origin effect on circadian rhythms and embryo size in hybrids

(**A-B**) Embryos of C24XL*er*C (**A**) and L*er*CXC24 (**B**) were dissected 10 days after pollination (DAP) and cultured in media for one day (day 1). After four days (day 4) in culture, C24XL*er*C (**A**) and L*er*CXC24 (**B**) embryos were subjected to bioluminescence assays. Scale bar = 0.1 mm. (**C**) Mean values (± s.e.m.) of bioluminescence counts in cultured embryos of the reciprocal F1 hybrids L*er*CXC24 (red) and C24XL*er*C (blue).

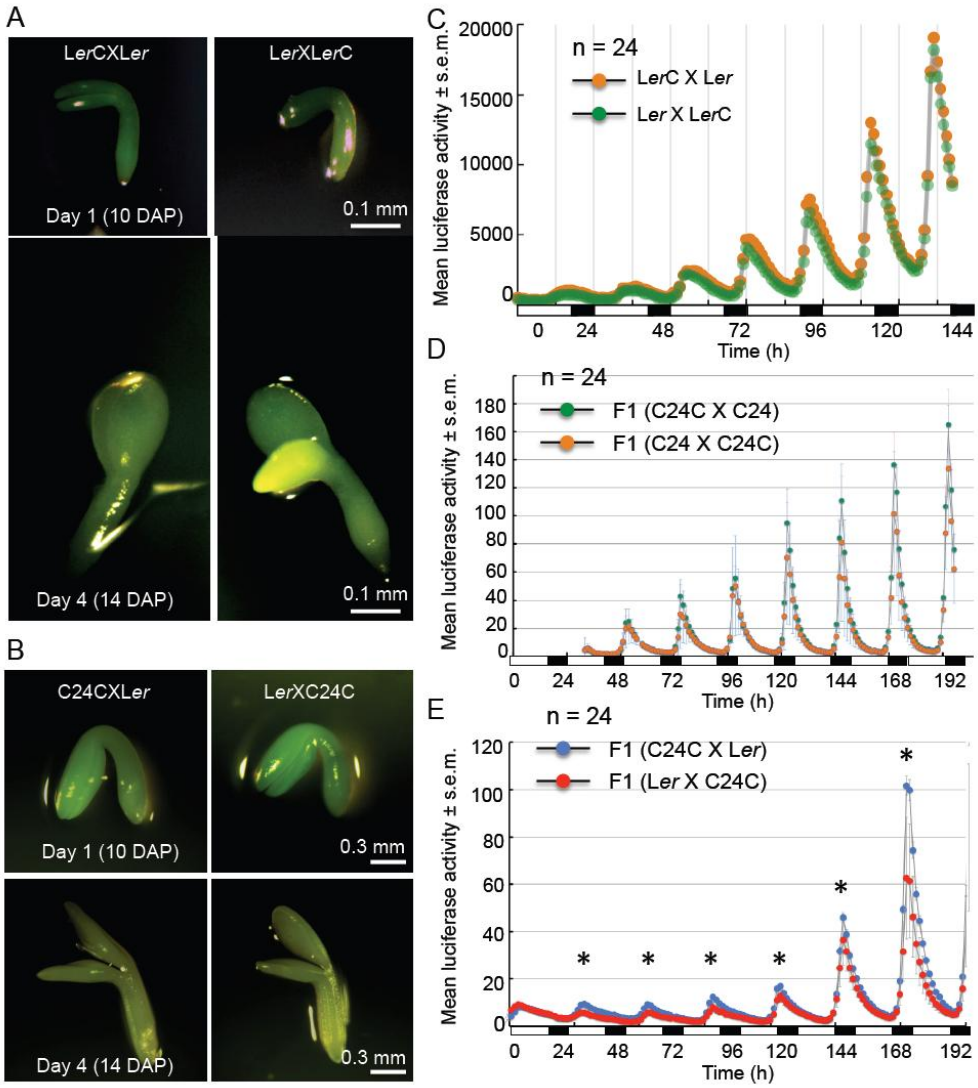


Figure 3.8 Parent-of-origin effects of *proCCA1:LUC* expression in embryos of reciprocal hybrids

(**A-B**) Embryos of L*er*XL*er*C (left panel) and L*er*CXL*er* (right panel) (**A**) and C24CXL*er* (left panel) and L*er*XC24C (right panel) (**B**) were dissected 10 days after pollination (DAP) and cultured in the medium for one day (day 1, upper panel) or four days (day 4, lower panel), when the embryos were subjected to bioluminescence assays. The scales are 0.1 mm (**A**) and 0.3 mm (**B**). (**C**) Mean values (± s.e.m.) of bioluminescence counts (Y-axis) in cultured embryos of L*er*XL*er*C (red) and L*er*CXL*er* (blue) reciprocal hybrids. (**D**) Bioluminescence counts (Y-axis) in cultured embryos of the reciprocal hybrids C24(*proCCA1:LUC*) or C24CXC24 (green) and C24XC24C (orange). (**E**) Mean values (± s.e.m.) of bioluminescence counts (Y-axis) in cultured embryos of the reciprocal hybrids C24CXL*er* (blue) and L*er*XC24C (red). Asterisks indicate statistical significance ($p < 0.05$).

A model for parent-of-origin effects on altered circadian rhythms and growth vigor in hybrids

The available data support a model that explains how changes in *CCA1* expression mediate growth vigor in hybrids (Figure 3.9). Changes in circadian expression amplitude (or phase) without altering the clock period can have significant consequences for clock-controlled metabolic rhythms in animals [152] as well as in plants [63]. Epigenetic repression of the maternal *CCA1* allele is correlated with increased starch content and biomass in *Arabidopsis* hybrids and allotetraploids [63]. *CCA1* repression is correlated with growth vigor in *A. thaliana* hybrids [31, 75], with repression of *CCA1* expression established during embryo development, which requires *AGO4* [133, 136]. Quantitative variation of CHH methylation levels in a regulatory motif (G-box and CHE-binding site, GTBS) containing region of the *CCA1* promoter region correlates with *CCA1* expression and depends on the AGO4-mediated pathway, possibly through interactions with 24-nt siRNAs. There is evidence for maternal transmission of 24-nt siRNAs in endosperm [153, 154] and expression changes of 24-nt siRNAs in *A. thaliana* hybrids relative to the parents [13, 31]. Collectively, these data suggest a role for siRNAs in the parent-of-origin effect on *CCA1* expression. AGO4 could recognize maternal and paternal siRNAs and guide parent-of-origin effects on DNA methylation. This discrimination mechanism may not solely depend on the primary DNA sequence because promoter sequences between two ecotypes are the same. The promoters could be associated with differential modifications of chromatin including histone acetylation and methylation. For example, H3K27me3 could induce CHG methylation through action of CMT3 [155] or CHH methylation through the action of CMT2 [156]. As a result, methylation of the *CCA1* promoter region inhibits binding of *CHE* and other proteins to GTBS, altering *CCA1* expression. When the CHH methylation level in the promoter is

high, *CCA1* is repressed and biomass is increased (Figure 3.9, right). Disruption of *AGO4* in the F1 crosses reverses CHH methylation levels in the promoter of paternal and maternal *CCA1* alleles and their expression directions, leading to altered biomass accumulation (Figure 3.9, bottom). When CHH methylation levels in the promoter are reduced, the paternal *CCA1* is increased, and biomass is also decreased (Figure 3.9, bottom left). *AGO4-1* is required for changes in CHH methylation but not sufficient to alter biomass vigor at the statistically significant level ($p = 0.052$). This is probably because biomass vigor is affected by many other epigenetic factors and metabolic pathways [14]. For example, disruption of RdDM pathways could alter expression of other genes in stress response and metabolic pathways, which can also influence growth vigor. The opposite is true in the reciprocal cross (Figure 3.9, bottom right). However, disruption of maintenance of DNA methylation (mainly CG) through *ddm1* or *met1* mutations does not change the parent-of-origin effect on *CCA1* expression and biomass.

Methylation and small RNA changes were previously observed in the same hybrids between C24 and L*er* [13, 31], but correlation of *CCA1* and *LHY* repression with specific methylation sites was not obvious. This is probably because genome-wide assays could not provide in-depth analysis of these loci in specific regions because the read coverage ranges from 2 to 47% of the genome [31]. We predict that siRNA and RdDM pathways may also affect other targets or regulators such as *LHY*, *TOC1*, and *CHE* in the circadian feedback loop and related networks, which in turn can directly or indirectly mediate *CCA1* expression [68, 157, 158]. Reduction in *CCA1* expression peaks or transcript levels promotes expression of downstream genes that are negatively regulated by *CCA1* abundance, as shown in *Arabidopsis* diploids and allopolyploids [63]. This altered circadian regulation could affect photosynthetic and metabolic pathways that are

altered in F1 hybrids [30, 49], as well as overall regulatory networks related to growth and development [22].

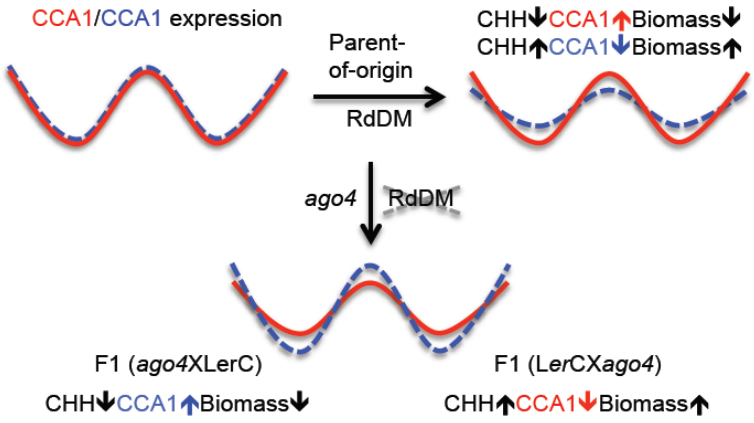


Figure 3.9 A model for the parent-of-origin effect on circadian rhythms and growth vigor in hybrids

Upper left: Maternal (red, solid line) and paternal (blue, dashed line) *CCA1* alleles are equally expressed in reciprocal crosses in the same genotype. Upper right: In reciprocal F1 hybrids, CHH methylation levels are low in the promoter of the maternal *CCA1* allele, and its expression is high; whereas CHH methylation levels are high in the promoter of the paternal *CCA1* allele, and its expression level is low. This parent-of-origin effect on *CCA1* expression is anti-correlated with biomass accumulation. Bottom: In F1 crosses involving the *ago4* mutant, disruption of RdDM leads to lower CHH methylation levels in the promoter of the paternal allele than that of the maternal allele. As a result, paternal *CCA1* expression is decreased, and biomass is increased (left). In the reciprocal cross, maternal *CCA1* expression is lower than that of the paternal allele (right). The reversal of the parent-of-origin effect on *CCA1* expression leads to increased levels of biomass accumulation.

The parent-of-origin effect on circadian rhythms and growth vigor is consistent with the parental conflict theory for imprinting in mammals and flowering plants, which predicts that the maternal genome provides factors inhibiting growth, whereas the paternal genome carries the factors promoting growth [126, 146, 147, 159]. This parental

conflict theory could apply to the maternal effect of the clock function on growth vigor during early stages of embryo development in hybrids and sexually reproducing organisms. When the maternal *CCA1* is repressed, growth vigor is increased. When the maternal *CCA1* expression is upregulated, growth vigor is reduced. This parent-of-origin effect on circadian rhythms and growth vigor in embryos is likely a general phenomenon and trans-generational in plant hybrids and allopolyploids and possibly in sexually reproducing organisms including mammals.

Chapter 4. Natural variation and timing of stress responses promotes heterosis in hybrids

Background and rationale

Emerging genomic and epigenetic data suggest that heterosis arises from allelic interactions between parental genomes, leading to changes in gene networks that promote growth vigor in hybrids [14]. Changes in DNA methylation and small RNAs have been associated with F1 hybrids in *A. thaliana* [13, 31], rice [29, 57], and maize [61]. Moreover, circadian-mediated regulatory networks, which promote growth and fitness in plants and animals [69, 70], control output pathways associated with heterosis [63]. In *Arabidopsis* allotetraploids and *A. thaliana* F1 hybrids, epigenetic regulation of circadian oscillator genes leads to altered circadian rhythms, which contributes to heterosis through photosynthetic and metabolic pathways [63]. Stress-responsive genes, which are regulated by the circadian clock [77-81], are also repressed in *Arabidopsis* allotetraploids [32].

The available data imply effects of altered circadian rhythms and stress responses on growth vigor in hybrids, but the underlying mechanisms are unknown. Here both genome-wide and transgenic approaches were employed to identify additional circadian-regulated candidate genes with novel roles in promoting biomass heterosis. Using mRNA-seq to analyze gene expression levels at 3 time points, it was found that hybrids differentially express stress and circadian-regulated genes in a time of day-specific manner. Circadian clock regulators, including *CCA1* and/or *LHY*, were shown to mediate natural variation of expression rhythms of stress-responsive genes such as *ACD6* and *COR78* among different ecotypes, which could be used to predict heterosis. The parents with larger expression differences between stress-responsive genes often produced

higher-vigor hybrids, while those with smaller expression differences made lower-vigor hybrids. Some stress-responsive genes were repressed in hybrids under normal conditions but induced under stress conditions at certain times of the day, balancing the tradeoff between heterosis and stress responses. In addition, repressing or overexpressing *ACD6* or *COR78* altered biomass accumulation. Lastly, it was found that methylation levels in hybrids may contribute to the genome-wide downregulation of circadian-regulated stress-responsive genes. Results from this work provide advances in the mechanistic understanding of heterosis, as well as effective selection criteria for parents to be used for producing high-yield hybrids.

Materials and methods

Plant growth and materials

Plant materials included the following *Arabidopsis* thaliana ecotypes, which were used to generate F1 hybrids; C24, Columbia (Col), Cvi-0 (CS22614), Est-1 (CS22629), L*er*, Nd-1 (CS22619), Oy-0 (CS22658), Sorbo (CS22653), Wei-0 (CS22622), and Ws. Crossing was carried out as previously described [75]. *acd6-1* was a generous gift from Dr. Detlef Weigel (Max Planck Institute for Developmental Biology, Tübingen, Germany) and *cca1-11* (CS9378), *lhy21* (CS9379), and *cca1-11lhy21* (CS9380) T-DNA insertion mutants in the Ws background were obtained from the Arabidopsis Biological Resource Center (ABRC). All plants were grown under a 16/8-h light/dark cycle with temperatures of 22°C (light) and 20°C (dark) on soil, and rosette leaves from ~3 week old plants before flowering were collected for RNA analysis, unless otherwise noted. For plant transformation, 4-5 week old plants were used for *Agrobacterium tumefaciens*–mediated transformation through floral dipping [143].

Biomass measurement

Biomass was measured as previously described [75], and the number of plants per replicate is indicated in Figure legends.

Trypan blue staining

Trypan blue staining was performed as previously described [160], except that 20 mg of Trypan blue was added to the lactophenol solution and leaves were destained in chloral hydrate solution for 30 minutes.

mRNA-seq library preparation

Total RNA was isolated from mature leaves (3 week-old *A. thaliana* plants) before bolting using plant RNA purification reagent (Invitrogen). The integrity of extracted RNA was analyzed by resolving 1 μg RNA in a 1%-formaldehyde denaturing agarose gel. An aliquot of 10 μg total RNA was used to prepare mRNA-seq libraries using Illumina mRNA-seq sample preparation kit (Illumina). Libraries with a range of 200±25 bp were excised and amplified for subsequent cluster generation and massively parallel sequencing using the Illumina Genome Analyzer GII with read lengths of 85 bp.

mRNA-seq read mapping and expression quantification

mRNA-seq reads from 12 libraries (Col, C24, ColxC24 and C24xCol at three time points) were mapped to the TAIR9 genome and cDNA sequence using BFAST (http://bfast.sourceforge.net) [161]. Transcript levels were quantified by counting RPKM as previously described [162]. The RNA-sequence data has been deposited at GEO (http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE51578.

Bioinformatics analyses

Differentially-expressed genes were identified using the R-package DEGseq with a $\log_2$ fold change > |0.5| [163] and q-value < 0.05. Heat maps were generated using

Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/). Circadian-regulated genes were identified by comparing a given dataset (i.e. DEG or MPV gene lists in parents and hybrids, respectively) with the intersection of two different microarray experiments (Covington + Edwards intersection [67]). IGV was used to display publically available mRNA-seq, sRNA-seq, and methylation data from Shen *et al* (GSE34658) [31].

RNA Extraction, cDNA synthesis, and qRT-PCR

Tissue collected for gene expression analysis was collected from plants before bolting (3 week old plants unless noted otherwise) at indicated Zeitgeber time (ZT0 = dawn) [68]. Total RNA was extracted using Concert Plant RNA Reagent (Invitrogen) and digested with RQ1 RNase-Free DNase (Promega) according to the manufacturer's instructions. cDNA was synthesized using 1μM oligo dT (12-18) primer (GeneLink) from 1 μg DNase-treated RNA using the Omniscript RT Kit (Qiagen) according to the manufacturer's instructions. For qRT-PCR, FastStart Universal SYBR Green Master (Rox) (Roche Applied Science) was used for PCR in the presence of gene-specific primers and 2 μL of diluted cDNA template. Expression levels of target genes were normalized against transcript levels of *ACT7* and relative expression levels of transcripts were calculated. The primer sequences were as follows: *ACT7-F*, 5'-GTCTGTGACAATGGAACTGGAA-3'; *ACT7-R*, 5'-CTTTCTGACCCATACCAACCAT-3'; *ACD6-F*, 5'-ATCACTGCAATTGCCCATGT-3'; *ACD6-R*, 5'-ACACGCCACACAACCAAAA-3'; *COR78-F*, 5'-CTTGATGGTCAACGGAAGGT-3'; *COR78-R*, 5'-CAATCTCCGGTACTCCTCCA-3'; *PR1-F*, 5'-CGTTCACATAATTCCCACGA-3'; *PR1-R*, 5'-AAGAGGCAACTGCAGACTCA-3'; *COR47-F*, 5'-CGGTACCAGTGTCGGAGAGT-

3’; *COR47-R*, 5’-ACAGCTGGTGAATCCTCTGC-3’; *PR5-F*, 5’-TCGAGGATTTTCAAGAACGC-3’; *PR5-R*, 5’-AAGCTTCGGTTTTTAAGGGC-3’; *COR15A-F*, 5’-GAAAAAAACAGTGAAACCGCAGAT-3’; *COR15A-R*, 5’-CCACATACGCCGCAGCTT-3’; *HSPRO2-F*, 5’-GAGGAAGACAGAGTGCGATAAG-3’; *HSPRO2-R*, 5’-CACTAACTGCCTATACCCAAAGA-3’; *RD22-F*, 5’-GATTCGTCTTCCTCTGATCTG-3’; *RD22-R*, 5’-TGGGTGTTAACGTTTACTCCG-3’; *PCC1-F*, 5’-ACAAATCTCACATCCTCACTCC-3’; *PCC1-R*, 5’-GCCCTGATGAAGTCTCTGAAG-3’; *RD28-F*, 5’-TTCGACGCAGAGGAGCTTACCA-3’; *RD28-R*, 5’-TACGAACTCGGCGATGACTGCT-3’; *LHCB1.4-F*, 5’-GCCTTCGCTACCAACTTCGTC-3’; *LHCB1.4-R*, 5’-AACCGGATACACACAACTCGATC-3’; *PORB-F*, 5’-GTGGACGGCAAGAAAACGTT-3’; *PORB-R*, 5’-GGCTCCAGTGACCACCACAT-3’; *CAB2-F*, 5’-ATTCGCAAGGAACCGTGAGCTAGA-3’; *CAB2-R*, 5’-TGAACCAGCCTTGAACCAAACTGC-3’; *LHCB4.2-F*, 5’-GCCGCCACTTCAACCGCCGCTGCTG-3’; *LHCB4.2-R*, 5’-CCCGTAGTCCCCGACAAGTGAACCG-3’; *CCA1-F*, 5’-CCTCGTCAGACACAGACTTCCA-3’; *CCA1-R*, 5’-CCGCAGTAGAATCAGCTCCAATA-3’; *LHY-F*, 5’-GGGACAAAGACTGCTGTTCAGAT-3’; *LHY-R*, 5’-TTTGTGAAGAACTTTTGTGCATGA-3’; *TOC1-F*, 5’-GTTGATGGATCGGGTTTCTC-3’; *TOC1-R*, 5’-TCATGACCCCATGCATACAG-3’.

Plasmid constructs

For luciferase reporter constructs, genomic DNA from Col, C24, and Ws was used to amplify *ACD6* and *COR78* promoter regions. SNPs between Col and C24 promoter and coding regions were identified using Polymorph (http://polymorph-clark20.weigelworld.org/index.php). The amplified fragments were cloned into pGEM-T vector (Promega) for sequence verification. The *promoter:LUC* plasmid constructs were generated by inserting luciferase gene between the restriction enzyme sites *Nco*I and *Bam*HI in the pFAMIR plasmid (from the lab of Dr. Ramin Yadegari, University of Arizona). For 35S-driven overexpression constructs, the transcribed region of *COR78* from Col or C24 was cloned into the pF35SE vector (A vector for 35S-driven gene expression, courtesy of Dr. Ramin Yadegari, University of Arizona) between *Rsr*II and *Aat*II restriction sites. Artificial miRNAs were designed using the WMD3 web app (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi) against a conserved region in Col and C24 and then amplified using the pRS300 vector as a template. The amplified fragments were cloned into pGEM-T vector (Promega) for sequence verification. The artificial miRNAs were then cloned into the pF35SE vector between *Rsr*II and *Aat*II restriction sites for *amiACD6* and between *Xma*I and *Aat*II restriction sites for *amiCOR78*. All constructs were individually cloned into *Agrobacterium* strain GV3101 for plant transformation and seeds were screened on 1% (w/v) agar with Murashige and Skoog (MS) media containing 7.5 mg/L phosphinothricin. The primer sequences used to make constructs were as follows: *Xho*I *pCOR78-F*, 5'-CCATCTCGAGAGATTTGGGGTTTTGCTTTTG-3'; *Nco*I *pCOR78-R*, 5'-CCATCCATGGGAGTAAAACAGAGGAGGGT-3'; *Xho*I *pACD6-F*, 5'-CCATCTCGAGAAGAGTTTGTAGCCTATTCAAAG-3'; *Nco*I *pACD6-R*, 5'-CCATCCATGGGGTTATCGAGAGGAGTGGTGGTC-3'; *Rsr*II *COR78-F*, 5'-

CCATCGGACCGATCAAACAGAGGAACCACC-3'; *Aat*II *COR78-R*, 5'-CCATGACGTCTTAAAGCTCCTTCTGCACCGG-3'; *amiCOR78-1*, 5'-GATATAGGTAACTTCGTTGTCACCTCTCTTTTGTATTCCA-3'; *amiCOR78-2*, 5'-AGGTGACAACGAAGTTACCTATATCAAAGAGAATCAATGA-3'; *amiCOR78-3*, 5'-AGGTAACAACGAAGTAACCTATTTCACAGGTCGTGATATG-3'; *amiCOR78-4*, 5'-GAAATAGGTTACTTCGTTGTTACCTACATATATATTCCTA-3'; *Aat*II *COR78-A*, 5'-CCATGACGTCCTGCAAGGCGATTAAGTTGGGTAAC-3'; *Xma*I *COR78-B*, 5'-CCATCCCGGGGCGGATAACAATTTCACACAGGAAACAG-3'; *amiACD6-1*, 5'-GATTAATGGTGACTAAAGGCCGTCTCTCTTTTGTATTCCA-3'; *amiACD6-2*, 5'-AGACGGCCTTTAGTCACCATTAATCAAAGAGAATCAATGA-3'; *amiACD6-3*, 5'-AGACAGCCTTTAGTCTCCATTATTCACAGGTCGTGATATG-3'; *amiACD6-4*, 5'-GAATAATGGAGACTAAAGGCTGTCTACATATATATTCCTA-3'; *Rsr*II *ACD6-A*, 5'-CCATCGGACCGCTGCAAGGCGATTAAGTTGGGTAAC-3'; *Aat*II *ACD6-B*, 5'-CCATGACGTCGCGGATAACAATTTCACACAGGAAACAG-3'.

Luciferase assays

Plants containing either *ACD6:LUC* or *COR78:LUC* constructs were analyzed using a Packard TopCount luminometer as previously described [164]. Seeds were sterilized with bleach and 75% ethanol and plated on 1% (w/v) agar with MS media containing 7.5 mg/L phosphinothricin. Seeds were stratified 2 days in the dark at 4ºC and then transferred into 16-h light and 8-h dark cycles for 7 days, and then transferred to MS containing no selection for 3 days. Seedlings were transferred to white microtiter plates (Nunc, Denmark) containing agar MS medium plus 30g sucrose/L and 30 μL of 0.5 mM luciferin (Gold Biotechnology). Microtiter plates were covered with clear plastic MicroAmp sealing film (Applied Biosystems, Foster City, CA) in which holes were

placed above each well for seedling gas exchange. One day after addition of luciferin, plates were moved to the TopCount and interleaved with two clear plates to allow light diffusion to the seedlings. All luciferase data were analyzed using the Biological Rhythm Analysis Software System (BRASS) (www.amillar.org). All period estimates were performed on rhythms for 24–120 hours using fast Fourier Transform–nonlinear least squares (FFT–NLLS) analysis.

Regression analysis

Regression analysis was performed using JMP 10 (JMP®, Version 10. SAS Institute Inc., Cary, NC, 1989-2007) with the default parameters and all graphs were generated using JMP.

Cold- and SA-treatments

For cold-induction of gene expression (cold-shock), 6 two week old seedlings (3 seedlings were used per replicate) were removed from agar media and were placed into 5 ml culture tubes containing room-temperature liquid MS media with 3% sucrose 24 hours before cold-treatment [165]. Cold-treatment was performed by placing the tubes in ice or leaving at 22°C for the control. After 60 minutes, tubes containing seedlings were placed back into 22°C and then whole seedlings (excluding roots) were snap frozen in liquid nitrogen at the ZT times indicated in the Figures.

For SA-induction of gene expression, two-week old seedlings growing on agar media were sprayed with 1 mM SA (Sigma-Aldrich). After 60 minutes, whole seedlings (excluding roots) were snap frozen in liquid nitrogen for the first time point, and then at the subsequent ZT times indicated in the Figures.

For longer-term cold stress application, two-week old seedlings growing on agar media were placed into a 4°C growth chamber of the same light intensity as control

seedlings which were left at 22°C. Every 2-3 days, seedlings were photographed and rosette diameter was quantified using ImageJ [166]. After two weeks, seedlings were snap frozen in liquid nitrogen at ZT0, ZT9, and ZT15. A batch of seedlings was transferred to soil, placed at 22°C to allow for recovery, and rosette diameter was measured every 2-3 days. After one week biomass was measured.

For longer-term SA-treatment, two-week old seedlings growing on agar media were sprayed with 1mM SA or water (control). Rosette diameter was measured every 2-3 days as described above. After one week seedlings were transferred to soil, sprayed with 1 mM SA again, and rosette diameter was measured every 2-3 days. After one more week, seedlings were snap frozen in liquid nitrogen at ZT0, ZT9, and ZT15 and biomass was measured.

To calculate relative expression ratios (R.E.R), expression levels of target genes were first normalized against transcript levels of *ACT7* and then the induced value was divided by the control value (indicated by a dashed line in the figures).

Relative growth rate (RGR) was calculated according to $RGR = (\ln D_2 - \ln D_1)/t_2 - t_1$, where D_1 and D_2 are rosette diameters at time points t_1 and t_2, respectively [167].

Results

Altered expression rhythms of photosynthetic and stress-responsive genes in F1 hybrids

F1 hybrids between *A. thaliana* Col and C24 ecotypes displayed biomass heterosis [9, 75] (Figure 4.1A), which is likely associated with gene expression changes in the hybrids. To determine if differential expression of genes occurs in a time of day-specific manner in hybrids, mRNAs were sequenced at zeitgeber time (ZT) 0 (dawn), ZT6 (midday), and ZT15 (dusk) from rosette leaves of 3-week old seedlings when

heterosis was most obvious in these hybrids. All experiments were performed in 16-hour light and 8-hour dark cycles because heterosis naturally occurs in diurnal conditions.

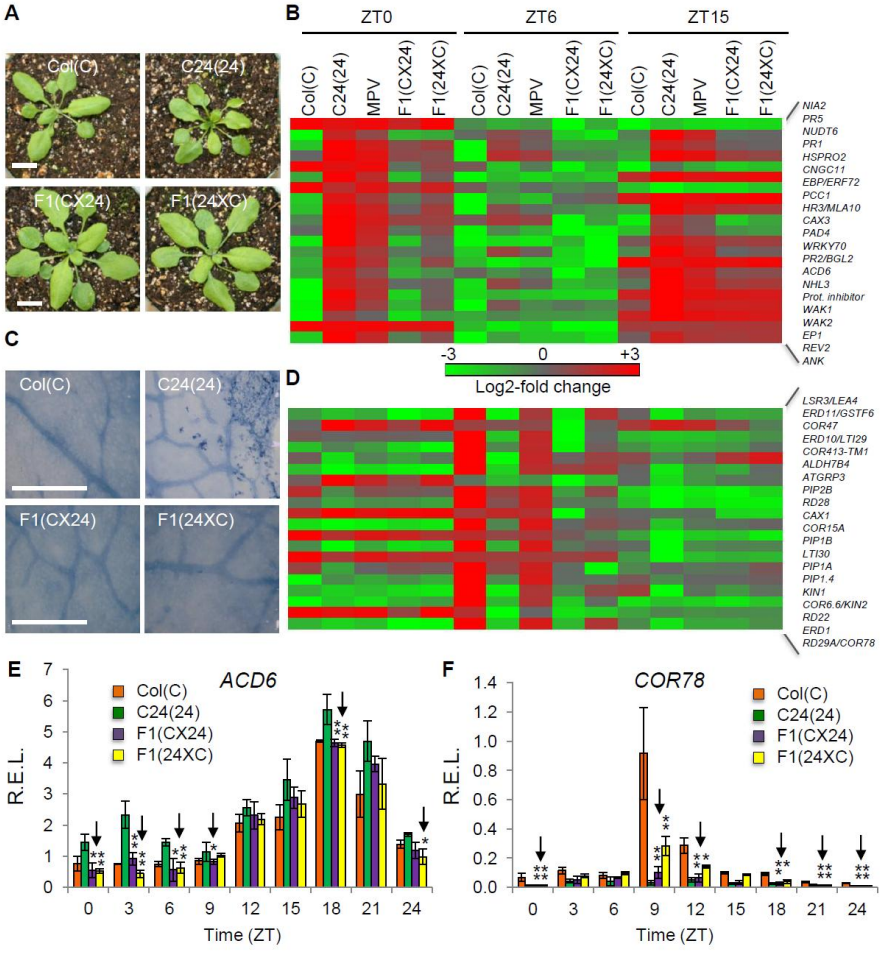


Figure 4.1 Expression of many abiotic and biotic stress-responsive genes is altered in reciprocal F1 hybrids relative to the parents

(**A, C**) Rosettes of 3-week-old plants (scale bars = 1 cm) (**A**) and the tenth leaf stained with Trypan blue (scale bars = 0.5 mm) (**C**). (**B, D**) Heatmaps of RNA-seq data showing subsets of genes in biotic (**B**) and abiotic (**D**) stress. Fold-change: from low (green) to high (red). (**E, F**) Relative expression levels (R.E.L.) of *ACD6* (**C**) and *COR78* (**D**) to those of *ACT7* every 3 hours in a 24-h period starting from dawn. Expression values were averaged from three biological replicates (error bars = s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, using two-tailed student's t-test (compared to the mid-parent value, MPV). Arrows indicate down-regulation in hybrids compared to the MPV.

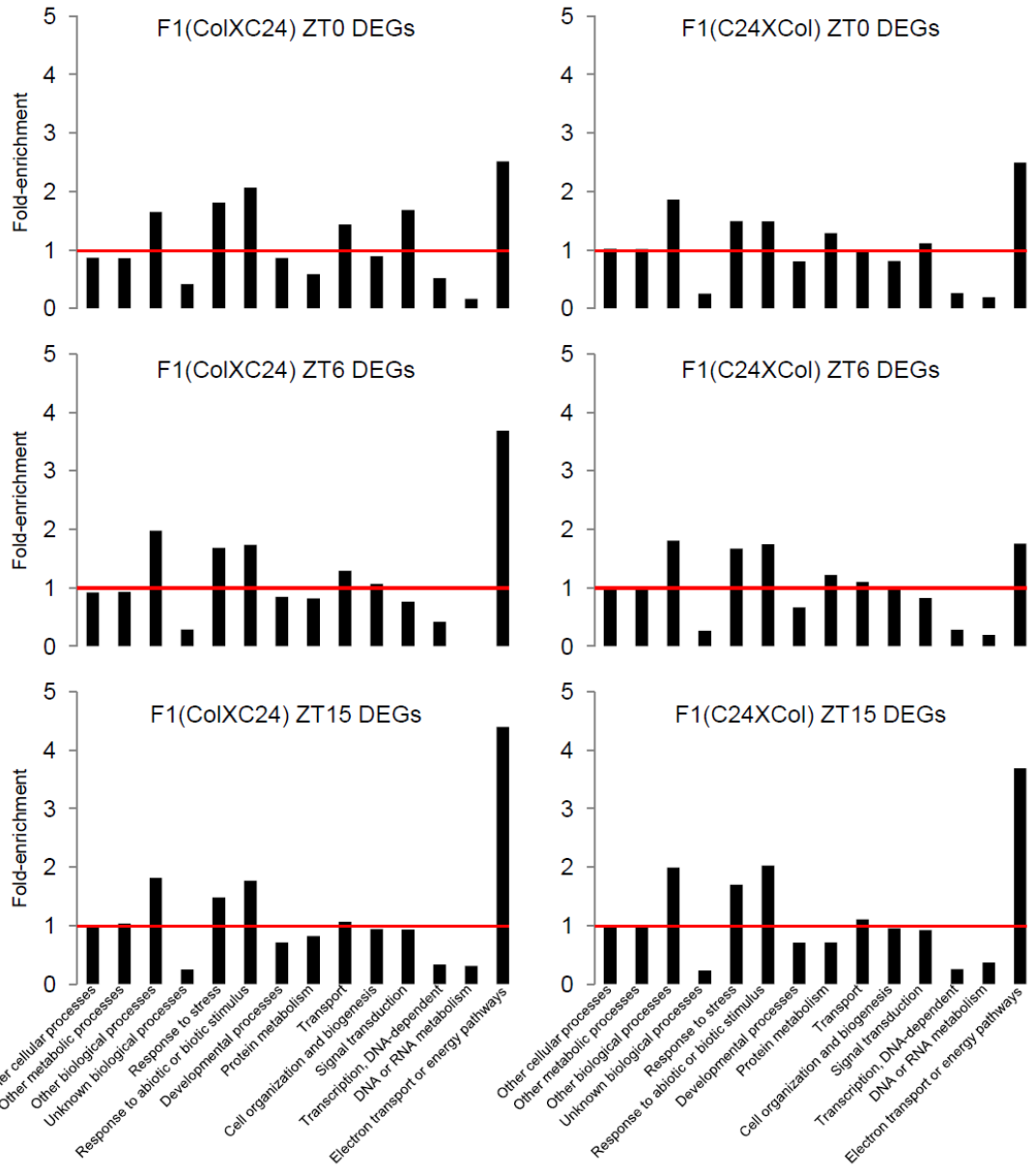


Figure 4.2 GO classifications of DEGs

Values above the red-line indicate enrichment compared to the whole genome. The GO classifications for DEGs were obtained from TAIR.

Differentially expressed genes (DEGs) between F1 hybrids and the mid-parent value (MPV) were estimated to be ~3% at ZT0, ~5% at ZT6, and ~1.3% at ZT15. Upon

examination of Gene Ontology (GO) categories of the DEGs, genes involved in stress responses and energy pathways were clearly differentially expressed at all time points, with no other GO terms being overrepresented (Figure 4.2). Therefore, genes involved in these categories were chosen for further examination. Over 30 genes involved in photosynthetic processes such as carbon fixation, light harvesting, and chlorophyll biosynthesis (Figure 4.3), including *CAB2*, *LHCB1.4*, *LHCB4.2*, and *PORB* (Figure 4.3B-E), were upregulated at ZT6, consistent with the observation that the diurnal upregulation of photosynthetic capacity enhances heterosis [30, 63, 75].

Interestingly, in the F1 hybrids many biotic stress-responsive genes were repressed in the morning (ZT0 and ZT6) (Figure 4.1B), while many abiotic stress-responsive genes were repressed in the afternoon (ZT6 and ZT15) (Figure 4.1D). RNA-seq data were validated by qRT-PCR analysis, and biotic stress-responsive genes such as *ACD6* were repressed most significantly from ZT0 to ZT6 (Figure 4.1E). In general, repression in hybrids was to the low-parent level. Similarly, *PATHOGEN AND CIRCADIAN CONTROLLED 1* (*PCC1*) and *HOMOLOG of SUGAR BEET HS1 PRO-2* (*HSPRO2*) were repressed in earlier parts of the day (Figure 4.4A,B). *PATHOGENESIS-RELATED GENES 1* and *5* (*PR1* and *PR5*) were repressed at all times and showed weak diurnal expression, except in the C24 ecotype (Figure 4.4C,D), even though both promoters contain CBS and EE motifs. Repression of abiotic (cold) stress-responsive genes, including *COLD REGULATED78*, *47*, and *15A* (*COR78*, *COR47*, and *COR15A*) (Figure 4.1F and Figure 4.4E,F) occurred in middle and later parts of the day. *RESPONSIVE TO DESSICATION 22* and *28* (*RD22* and *RD28*) were weakly repressed in the morning and afternoon, respectively (Figure 4.4G,H).

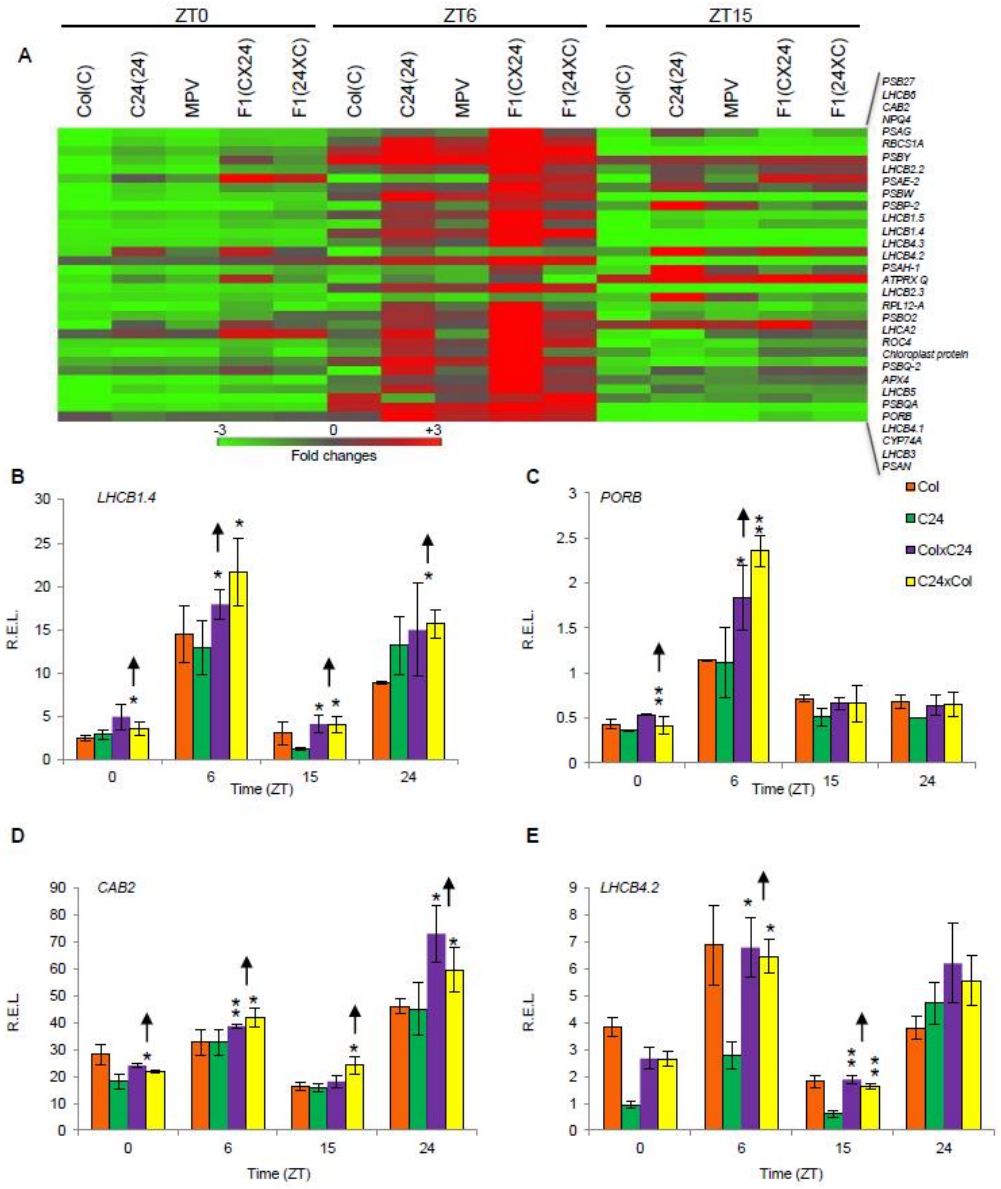


Figure 4.3 Photosynthetic genes are upregulated in F1 hybrids

(**A**) Heatmap of RNA-seq data showing a subset of genes in the photosynthetic pathway (ZT0 = dawn). Fold-change: Log2-transformed values from low (green) to high (red). R.E.L. of *LHCB1.4* (**B**), *PORB* (**C**), *CAB2* (**D**), and *LHCB4.2* (**E**) at 4 time points in reciprocal F1 hybrids between Col and C24, compared to the MPV. Values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, using two-tailed student's t-test (compared to MPV). Arrows indicate up-regulation in the hybrids compared to the MPV.

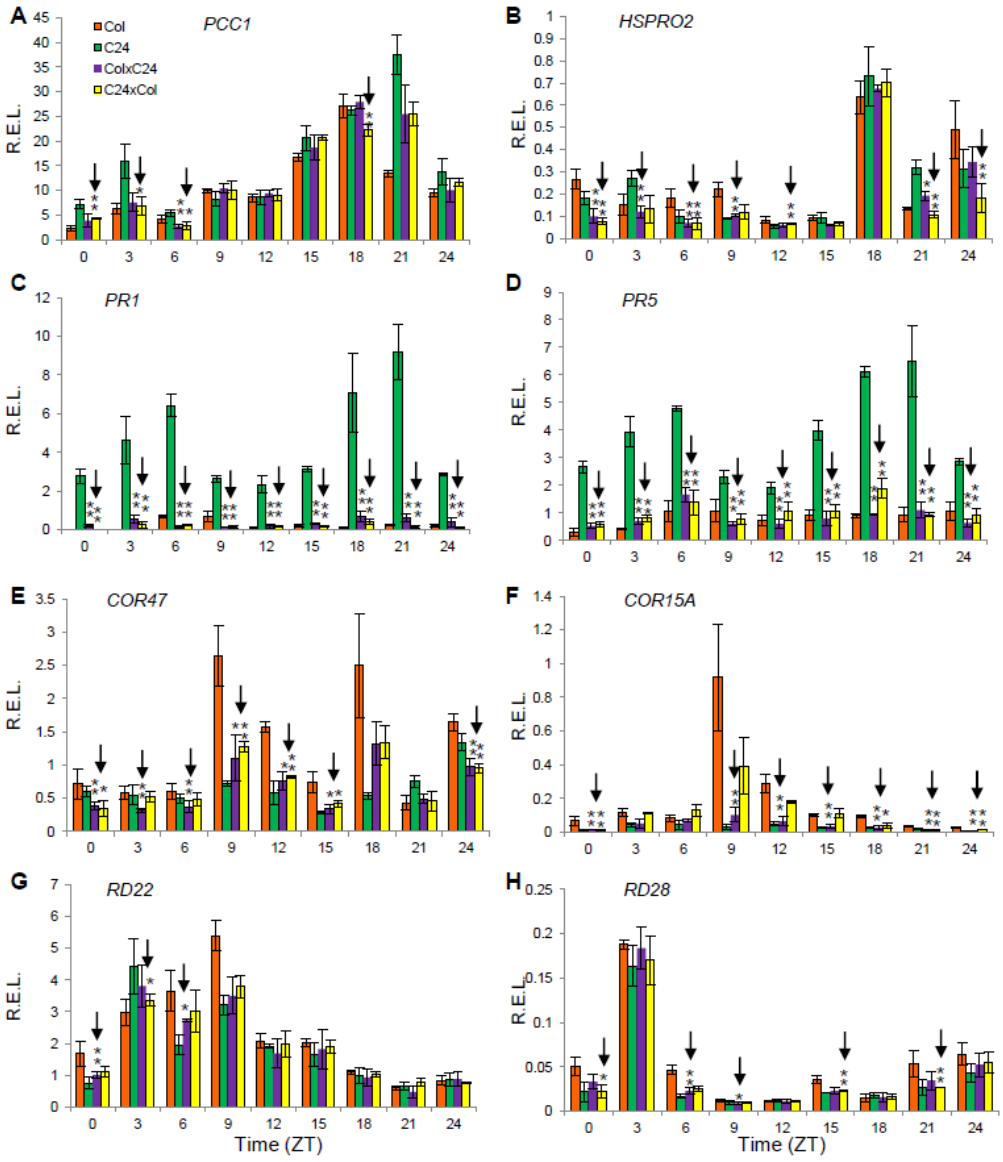
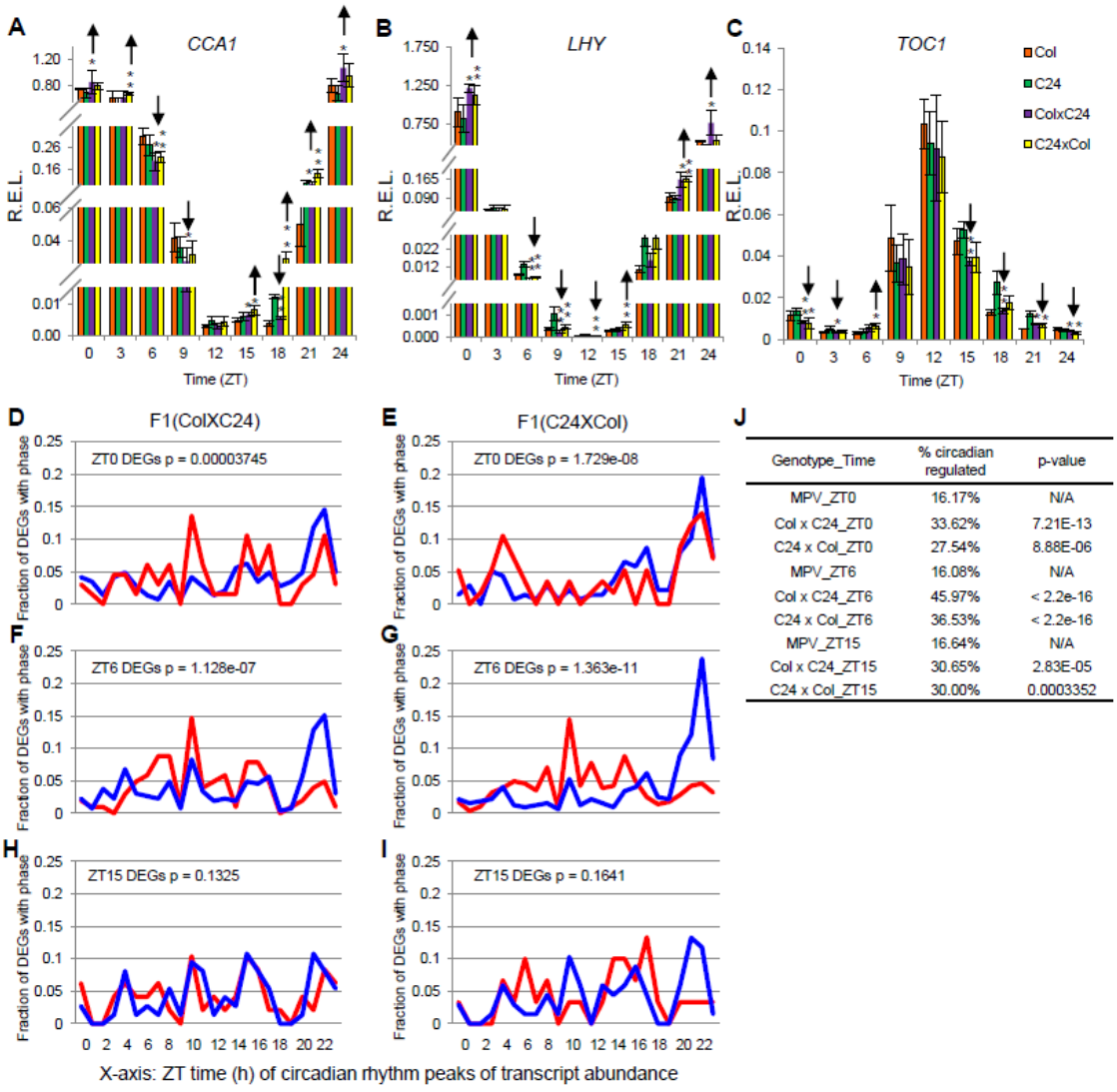


Figure 4.4 Biotic and abiotic stress genes show altered expression in F1 hybrids

R.E.L. of *PCC1* (**A**), *HSPRO2* (**B**), *PR1* (**C**), *PR5* (**D**), *COR47* (**E**), *COR15A* (**F**), *RD22* (**G**), and *RD28* (**H**) in a 24-h period starting from dawn. Values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, using two-tailed student's t-test (compared to the MPV). Arrows indicate down-regulation in hybrids compared to the MPV.



J

Genotype_Time	% circadian regulated	p-value
MPV_ZT0	16.17%	N/A
Col x C24_ZT0	33.62%	7.21E-13
C24 x Col_ZT0	27.54%	8.88E-06
MPV_ZT6	16.08%	N/A
Col x C24_ZT6	45.97%	< 2.2e-16
C24 x Col_ZT6	36.53%	< 2.2e-16
MPV_ZT15	16.64%	N/A
Col x C24_ZT15	30.65%	2.83E-05
C24 x Col_ZT15	30.00%	0.0003352

Figure 4.5 Perturbation of circadian genes in F1 hybrids

R.E.L. of *CCA1* (**A**), *LHY* (**B**), and *TOC1* (**C**) in a 24-h period starting from dawn in reciprocal F1 hybrids between C24 and Col. Values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, using two-tailed student's t-test (compared to the MPV). Upward and downward arrows indicate up- and down-regulation in the hybrids compared to the MPV. Phase distributions of differentially expressed genes (DEGs) in hybrids in ColxC24 (**D, F, H**) and C24xCol (**E, G, I**) at ZT0, ZT6, and ZT15. Blue and red lines indicate down- and up-regulated genes, respectively. Percentage of DEGs in hybrids which are circadian-regulated (**J**). p-values calculated using chi-square test of independence in (**I, J**).

Consistent with the diurnal regulation, DEGs were significantly enriched for evening element ($p = 0.018$) and CCA1-binding site (CBS) ($p < 10\text{-}6$) motifs, indicating that altered circadian rhythms mediate stress responses and photosynthesis in the hybrids. Indeed, circadian clock regulators *CCA1* and *LHY* were repressed in the middle of the day and upregulated around dawn (Figure 4.5A,B), and their feedback regulator *TOC1* showed opposite expression changes (Figure 4.5C). Altered expression of circadian regulators and slave oscillators, such as *TOC1*, *LHY*, and *GLYCINE RICH PROTEIN7* (*GRP7*), correlates with phase changes in downstream genes, as a result of the clock regulation [78, 168, 169]. Consistent with this observation, phases of down- and up-regulated DEGs at ZT0 and ZT6 in the hybrids were clustered towards dawn and in the middle and later parts of the day, respectively, although DEGs at ZT15 showed a phase distribution similar to the parents probably because the number of DEGs is low at ZT15 (Figure 4.5D-I). Overall, DEGs were statistically significantly enriched for circadian-regulated genes (Figure 4.5J). These data suggest that expression of circadian clock genes and stress-responsive genes is coordinately changed in the hybrids to promote growth.

Natural variation in stress-responsive gene expression as one predictor for heterosis

Stress responses are adaptive traits and vary among ecotypes. For example, Col (Columbia, Missouri) is more cold tolerant than C24 (Coimbra, Portugal) [11, 99, 100], whereas C24 has a higher level of salicylic acid (SA) and shows more necrosis on mature leaves (Figure 4.1C) [170, 171]. In general, this natural variation of biotic and abiotic stress responses is consistent with expression levels of biotic-stress responsive genes (Figure 4.1B and Figure 4.4A-D) and abiotic-stress responsive genes (Figure 4.1D and Figure 4.4E-H).

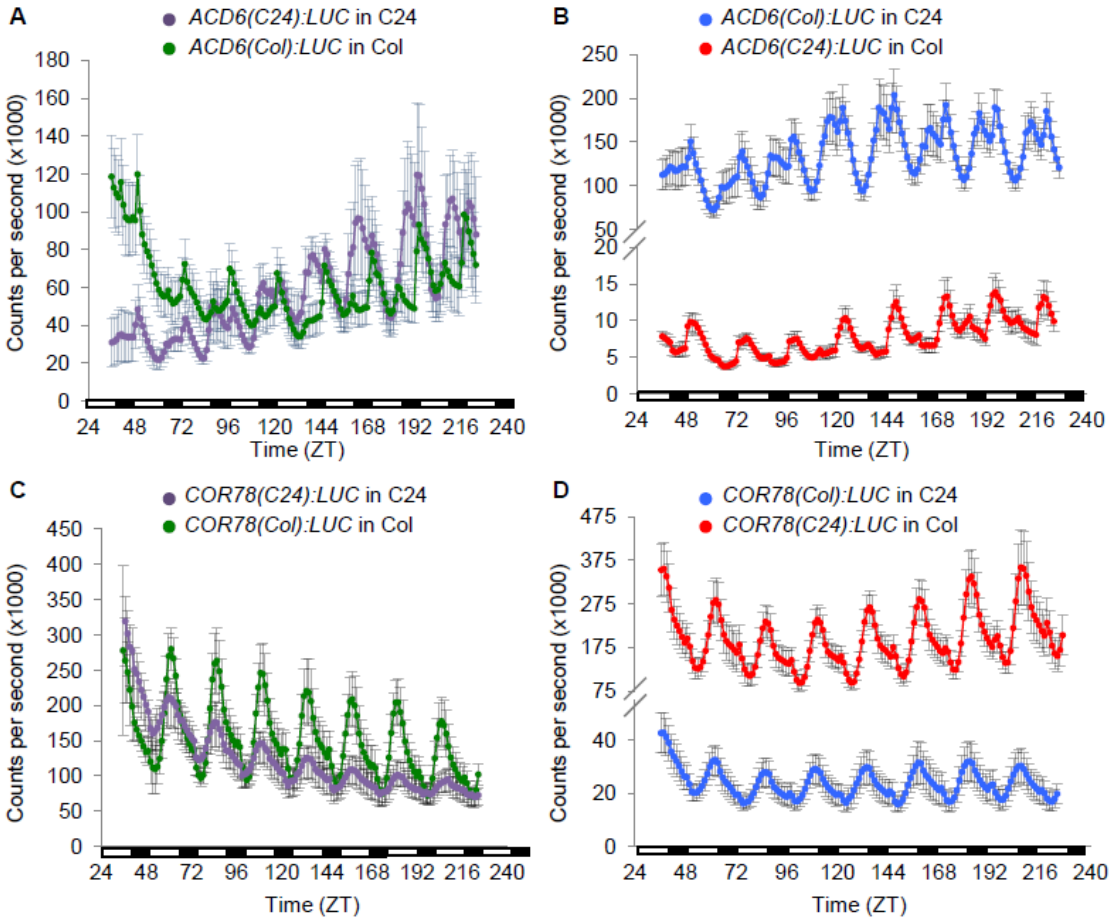


Figure 4.6 *ACD6* and *COR78* were subjected to diurnal regulation and differentially expressed between two ecotypes, Col and C24

(**A, B**) Bioluminescent activities (Y-axis, counts per second in 1,000) of 16 transgenic (T2) seedlings that expressed *ACD6(Col):LUC* in Col (green) or *ACD6(C24):LUC* in C24 (purple) (**A**) and *ACD6(Col):LUC* in C24 (blue) or *ACD6(C24):LUC* in Col (red) (**B**). (**C, D**) Bioluminescent activities of 16 transgenic seedlings that expressed *COR78(Col):LUC* in Col (green) or *COR78(C24):LUC* in C24 (purple) (**C**) and *COR78(Col):LUC* in C24 (blue) or *COR78(C24):LUC* in Col (red) (**B**). X-axis: hours (ZT0 = dawn) with nine diurnal cycles of light (open) and dark (filled). One of three replicate experiments was shown (error bars = s.e.m.).

Using *ACD6* and *COR78* as examples of biotic and abiotic stress-responsive genes, respectively, the natural variation between ecotypes was further characterized. Promoters of *ACD6* and *COR78* derived from Col and C24 were used to drive luciferase

(LUC) reporter expression. Four constructs, *ACD6*(C24):*LUC*, *ACD6*(Col):*LUC*, *COR78*(C24):*LUC*, and *COR78*(Col):*LUC*, were transformed into both Col and C24. Both *ACD6:LUC* and *COR78:LUC* displayed rhythmic activities, suggesting circadian regulation of these stress-responsive genes. Similar to the expression of endogenous genes (Figure 4.1E,F), *ACD6*(C24):*LUC* in C24 was expressed relatively higher than *ACD6*(Col):*LUC* in Col (Figure 4.6A), and *COR78*(Col):*LUC* in Col was expressed higher than *COR78*(C24):*LUC* in C24 (Figure 4.6C). The expression differences were amplified when they were expressed in reciprocal ecotypes. *ACD6*(Col):*LUC* expression amplitudes in C24 were 10-15-fold higher than *ACD6*(C24):*LUC* levels in Col (Figure 4.6B). Likewise, *COR78*(C24):*LUC* amplitudes in Col were 7-8 fold higher than *COR78*(Col):*LUC* in C24 (Figure 4.6D). Thus, divergence between genetic backgrounds and between promoter sequences (Figure 4.7A) may contribute to natural variation of stress-responsive genes in ecotypes that are differentially adapted to biotic and abiotic stresses.

To test a relationship between natural variation of stress responses and heterosis, expression levels of two biotic and two abiotic stress-responsive genes were surveyed at three time points in 10 ecotypes originating from diverse geographical locations [12] (Figure 4.7B-G). The resulting data were tested for correlation with biomass heterosis in hybrids. Among these ecotypes, expression variation of stress-responsive genes correlated with their geographical locations. For example, abiotic genes, including *COR78*, were poorly expressed in C24 and Cvi (Cape Verde Islands), but biotic genes, including *ACD6*, were highly expressed, probably because they are adapted to hot and dry conditions (Figure 4.7F,G) [171, 172]. However, in Col, L*er* (Germany), and Ws (Russia) that are adapted to a colder climate, *COR78* was highly expressed (Figure 4.7F), but *ACD6* and *PR1* were poorly expressed [98] (Figure 4.7D,E). Expression levels of

stress-responsive genes were less variable in other ecotypes, or showed similar patterns (Figure 4.7D-G). Interestingly, no ecotype had low expression levels of both biotic and abiotic stress genes, indicating that the trend of reduced expression of both types of stress genes is unique to the hybrid state.

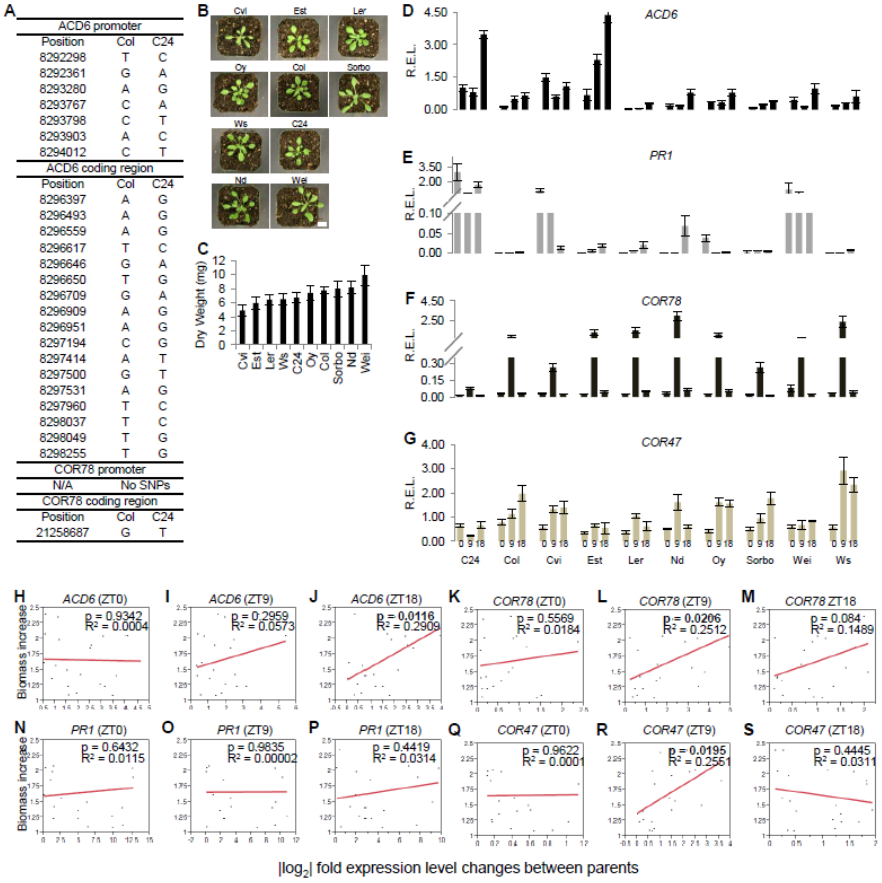


Figure 4.7 Natural variation of stress responsive gene expression at specific times of day is associated with biomass heterosis

(**A**) SNPs between Col and C24 ecotypes in *ACD6* and *COR78*. (**B**) Seedlings of 3-week-old plants in various ecotypes (scale bar = 1 cm). (**C**) Biomass (dry weight, mg) of the ecotypes in three replications (mean ± s.d.). (**D-G**) R.E.L of *ACD6* (**C**), *PR1* (**D**), *COR78* (**E**) and *COR47* (**F**) in ecotypes at ZT0, ZT9, and ZT18. Values are averages from three biological replicates (± s.d.). (**H-S**) Biomass increase relative to MPV (Y-axis) in various hybrids was plotted against absolute values of the log2-fold expression level changes (X-axis) in different stress genes at three time points (ZT0, 9, 18). Regression lines (red) were statistically significant for *ACD6* at ZT18 (**J**), *COR78* at ZT9 (**L**) and *COR47* at ZT9 (**R**).

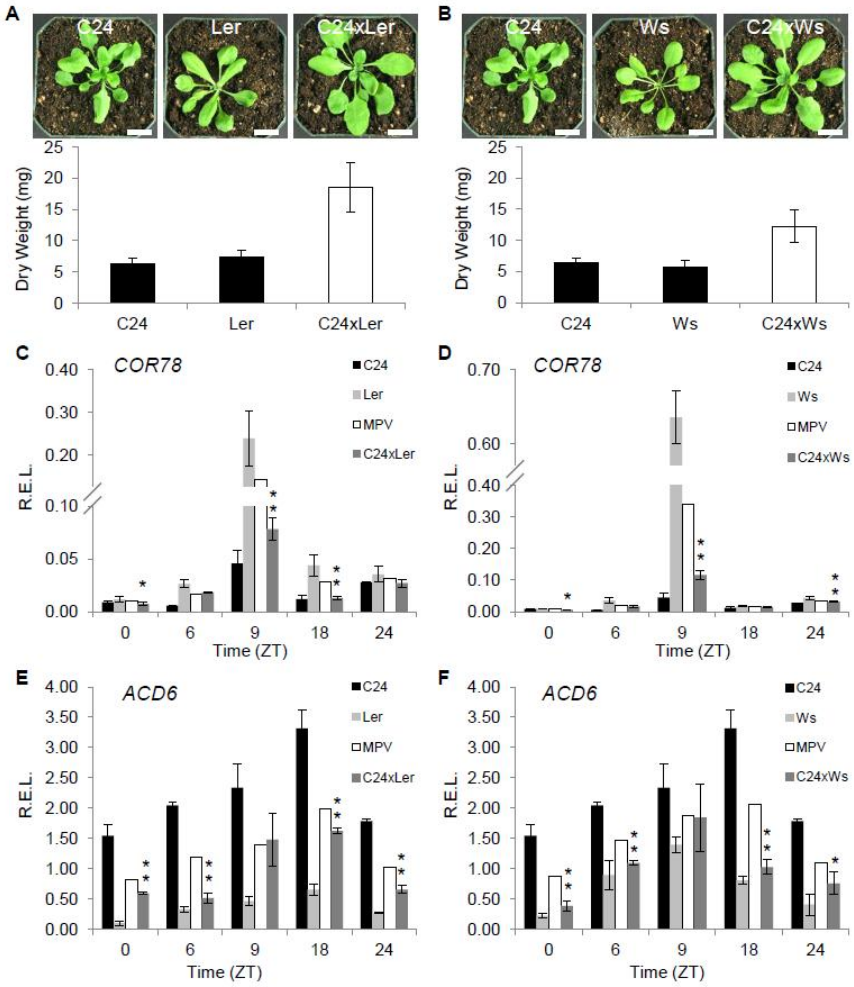


Figure 4.8 Stress-responsive gene expression in vigorous F1 hybrids

(**A**) Rosettes of 3-week-old plants (upper panel) and biomass (lower panel) in F1 and its parents C24 and L*er*. (**B**) Rosettes of 3-week-old plants (upper panel) and biomass (lower panel) in F1 and its parents C24 and Ws. Scale bars = 1 cm for all images in (**A-B**). (**C-D**) R.E.L of *COR78* in C24 and L*er* and their F1 (**C**) and in C24 and Ws and their F1 (**D**) at 5 time points, compared to the MPV. Values were averaged from three biological replicates (± s.d.). (**E-F**) R.E.L of *ACD6* in C24 and L*er* and their F1 (**E**) and in C24 and Ws and their F1 (F) at 5 time points, compared to MPV. Values were averaged from three biological replicates (± s.e.m.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, using two-tailed student's t-test (compared to MPV).

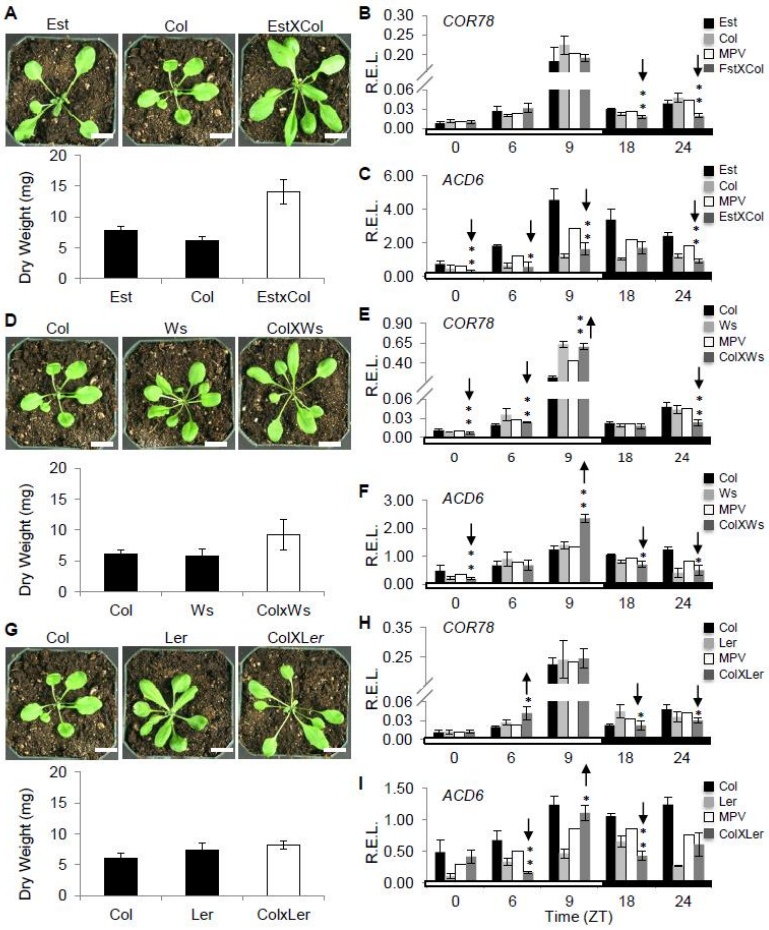


Figure 4.9 Stress-responsive gene expression in low-vigor F1 hybrids

(**A**) Seedlings of 3-week-old plants (upper, scale bars = 1 cm) and biomass (lower) in F1(EstXCol) and their parents Est and Col. (**B-C**) Relative expression levels (R.E.L.) of *COR78* (**B**) and *ACD6* (**C**) in F1(EstXCol) and their parents in a 24-hour light/dark cycle. (**D**) Seedlings of 3-week-old plants (upper) and biomass (lower) in F1(ColXWs) and their parents Col and Ws. (**E-F**) R.E.L. of *COR78* (**E**) and *ACD6* (**F**) in F1(ColXWs) and their parents in a 24-hour cycle. (**G**) Seedlings of 3-week-old plants (upper) and biomass (lower) in F1(ColXL*er*) and their parents Col and L*er*. (**H-I**) R.E.L. of *COR78* (**H**) and *ACD6* (**I**) in F1(ColXL*er*) and their parents in a 24-hour cycle. Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively.

Hybrids were made between the set of ecotypes, and heterosis was quantified as the mean increase in biomass relative to the MPV. This value was used as the dependent variable to perform regression analysis with expression level changes for a given stress-responsive gene between the parents as the independent and predicting variable at different time points (Figure 4.7H-S). Expression differences of the stress-responsive

genes between parental ecotypes at certain times of day, including *ACD6* at ZT18, *COR78* at ZT9, and *COR47* at ZT9 (Figure 4.7J, L, and R), were statistically significantly correlated with biomass heterosis. Therefore, the time of day when gene expression differences between parents are measured is crucial for accurately predicting the level of heterosis in the hybrids.

This natural variation in stress responsive-gene expression between parents could be an indicator of genetic distance to predict the degree of heterosis. According to the prediction, higher heterosis in F1 hybrids could be obtained between parents with larger expression differences in stress-responsive genes. For example, Col has high expression of an abiotic gene (*COR78*) and low expression of a biotic gene (*ACD6*), while C24 has low expression of *COR78* and high expression of *ACD6*. The expression of both *COR78* and *ACD6* was repressed in the hybrids (Figure 4.1B,D), which led to high heterosis (Figure 4.1A,C). Using this concept, two additional high-vigor F1 hybrids (C24XL*er* and C24XWs) were predicted (Figure 4.8A,B), which correlated with higher levels of expression differences in *ACD6* and *COR78* between the parents (C24 vs. L*er* and C24 vs. Ws) (Figure 4.7C-F). On the contrary, if the parents (Col vs. Ws and Col vs. L*er*) had similar levels of *ACD6* and *COR78* expression (Figure 4.E,F,H,I), relatively low levels of heterosis were obtained in their F1 hybrids (ColXWs and ColXL*er*) (Figure 4.9D,G), as reported [9, 30, 75]. Moreover, in low-vigor hybrids, *COR78* and *ACD6* were upregulated at certain times of day, although they were repressed at other times; whereas in the high-vigor hybrids, both *COR78* and *ACD6* were overall repressed. Another F1 hybrid (EstXCol) displayed higher levels of biomass heterosis (Figure 4.9A) possibly because *ACD6* was expressed much higher in Est than in Col [98], although *COR78* expression levels were more similar. Both *ACD6* and *COR78* were repressed in these F1 hybrids (Figure 4.9B,C). In all hybrids tested, *CCA1* expression was repressed, and *TOC1*

expression was up-regulated (Figures 4.10 and 4.11) at certain times of day, linking circadian regulation of output stress-responsive genes with heterosis [31, 63, 75].

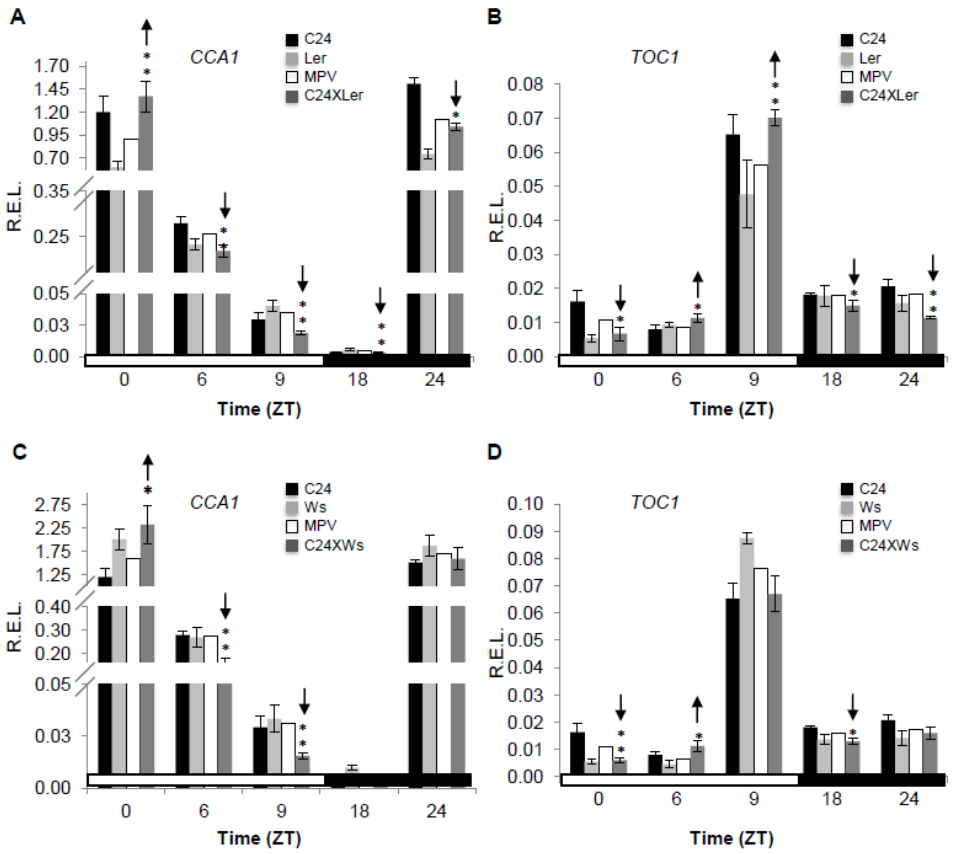


Figure 4.10 Altered expression of *CCA1* and *TOC1* in F1 hybrids with high levels of heterosis

(**A-B**) R.E.L. of *CCA1* (**A**) and *TOC1* (**B**) in F1(C24XL*er*) and its parents, C24 and L*er*, in a diurnal 24-hour cycle (light/dark). (**C-D**) R.E.L of *CCA1* (**C**) and *TOC1* (**D**) in F1(C24XWs) and its parents, C24 and Ws, in a diurnal 24-hour cycle (light/dark). Values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, compared to the MPV. Upward and downward arrows indicate up- and down-regulation in the hybrids compared to the MPV.

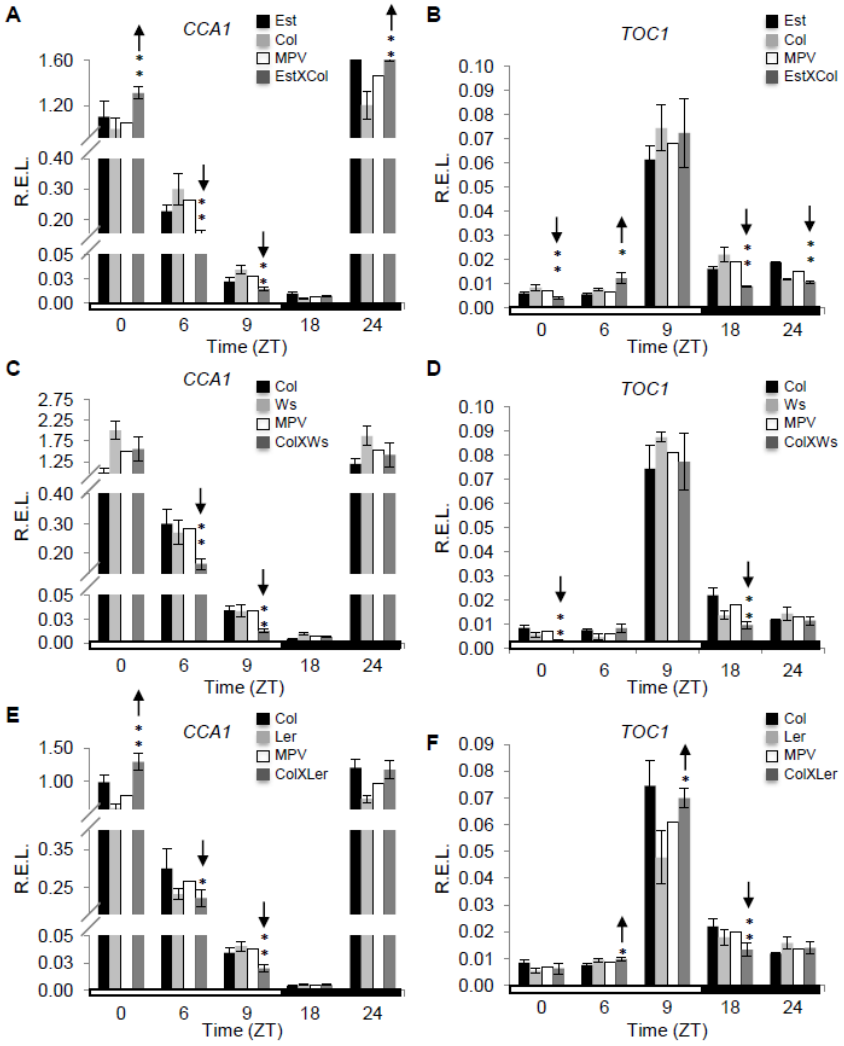


Figure 4.11 Altered expression of *CCA1* and *TOC1* in hybrids with low levels of heterosis

(**A-B**) R.E.L. of *CCA1* (**A**) and *TOC1* (**B**) in a 24-hour light/dark cycle in F1(EstXCol) and their parents, Est and Col. (**C-D**) R.E.L. of *CCA1* (**C**) and *TOC1* (**D**) in a 24-hour light/dark cycle in F1(ColXWs) and their parents, Col and Ws. (**E-F**) R.E.L. of *CCA1* (**E**) and *TOC1* (**F**) in a 24-hour light/dark cycle in F1(ColXL*er*) and their parents, Col and L*er*. Expression values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$, respectively, compared to the MPV. Upward and downward arrows indicate up- and down-regulation in the hybrids compared to the MPV.

Circadian clock regulates rhythmic expression of stress-responsive genes

The above data suggest that the clock regulates expression of stress-responsive genes and consequently heterosis. Circadian effects on stress-responsive gene expression were examined in transgenic plants that expressed *ACD6:LUC* and *COR78:LUC* in the wild-type (Ws) and *cca1-11*, *lhy21*, and *cca1-11lhy21* mutants. *ACD6:LUC* expression was inducible in response to SA treatment, a plant hormone which triggers inducible gene expression in response to biotic stresses [173] (Figure 4.12A). Likewise, *COR78:LUC* expression was inducible after 1 hour of cold-shock [165] (Figure 4.12B). Notably, *ACD6:LUC* expression became arrhythmic after SA application, whereas induction of *COR78:LUC* was rhythmic and remained elevated for ~36 hours before reaching the control level. In the absence of stress, the period of *ACD6:LUC* activity (WT period 24.26 ± 0.30 s.e.m.) was not significantly altered in the *cca1-11* (24.15 ± 0.22) or *lhy21* (24.49 ± 0.31) mutants, but was shortened ~1 hour in the *cca1-11lhy21* double mutant (23.17 ± 0.41, $p < 0.05$ two-tailed student's t-test) (Figure 4.12C). Expression amplitudes were significantly lowered in both *cca1-11* and *cca1-11lhy21* mutants. Both expression amplitudes and periods of *COR78:LUC* (WT 24.75 ± 0.11) were severely altered in single (*cca1-11*, 23.73 ± 0.51; *lhy21* 24.14 ± 0.30, $p < 0.05$) and double (22.61 ± 0.66, $p < 0.05$) mutants (Figure 4.12E), consistent with a previous study [77]. During stress conditions, *ACD6:LUC* was inducible in the single mutants as in the wild-type, but not in the double mutant (Figure 4.12D). Induction of *COR78:LUC* was compromised in all mutants tested (Figure 4.12E,F). These data suggest that circadian regulators, including *CCA1* and/or *LHY*, mediate the expression amplitudes and periods of both biotic (*ACD6*) and abiotic (*COR78*) stress-responsive genes during stress and non-stress conditions.

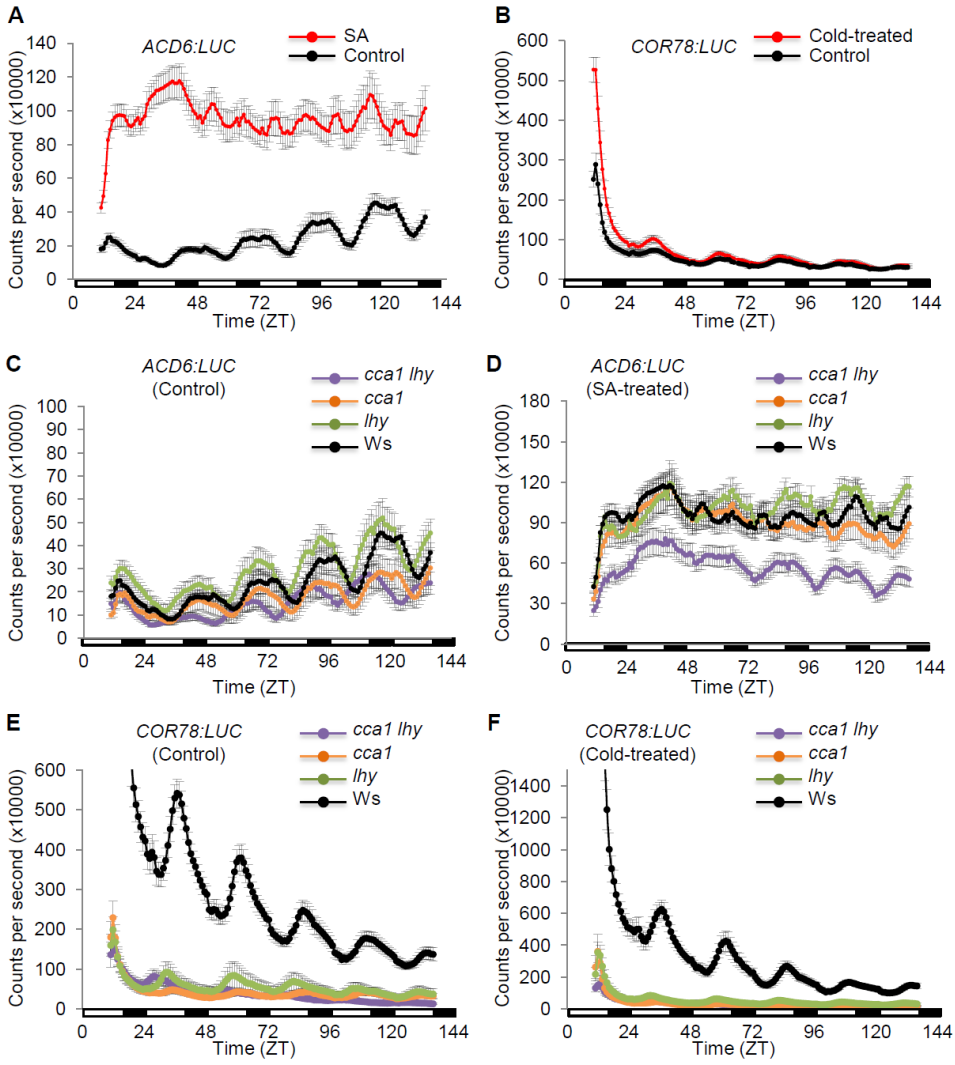


Figure 4.12 Rhythmic expression patterns of *ACD6* and *COR78* are mediated by components of the circadian clock

(**A, B**) Bioluminescent activities of *pACD6*(Ws):*LUC* in T3 transgenic seedlings that were treated with salicylic acid (SA, red) and control (black) (**A**) and of *pCOR78*(Ws):*LUC* in T3 transgenic seedlings that were treated with cold (red) and control (black) (**B**). (**C, D**) Altered bioluminescent activities of *pACD6*(Ws):*LUC* in T3 transgenic seedlings in *cca1-11* (orange), *lhy21* (green), *cca1-11 lhy21* (purple), and Ws (black) backgrounds, which were sprayed with water (control, c) or SA (**D**). (**E, F**) Altered bioluminescent activities of *pCOR78*(Ws):*LUC* in T3 transgenic seedlings in *cca1-11* (orange), *lhy21* (green), *cca1-11 lhy21* (purple), and Ws (black) backgrounds, which were subjected to control (**E**) or cold (**F**) treatments. Y-axis, Bioluminescent activities (counts per second in 10,000); X-axis: hours (ZT0 = dawn) with six diurnal cycles of light (open) and dark (filled). One of three replicate experiments (± s.e.m., n = 16 in each replicate).

Effects of repressing and overexpressing stress-responsive genes on biomass

There are tradeoffs between fitness and constitutive and induced expression of stress-responsive genes [93, 95, 167, 174]. Constitutive defense responses decrease biomass, growth rate, and seed production [167, 174], and R-genes reduce fitness in the absence of pathogens [93, 95, 175]. If altered expression of stress-responsive genes affects heterosis, knockdown and overexpressing stress-responsive genes (*ACD6* and *COR78*) in Col and C24 ecotypes would change biomass accumulation (Figure 4.13). Indeed, the dominant negative mutant *acd6-1* accumulated less biomass than the wild-type due to the production of higher amounts of SA and spontaneous cell death [98, 176] (Figure 4.13A). Conversely, knockdown of *ACD6* (*amiACD6*) using artificial microRNAs (amiRNA) showed a modest increase in biomass relative to Col (Figure 4.13B,D), whereas *amiACD6* in C24 increased biomass at statistically significant levels (Figure 4.13C,D). Knockdown of *COR78* also caused modest increases in biomass (Figure 4.13B,C,D). Overexpressing *COR78* in either Col or C24 caused biomass reductions at statistically significant levels (Figure 4.13B,C,E), although the reductions were not as severe as transgenic plants overexpressing the upstream regulator *C-REPEAT BINDING FACTOR* (*CBF*) genes, which cause overexpression of many downstream *COR* genes [94, 177]. Biomass increase or reduction in individual transgenic lines was anti-correlated with expression levels of *ACD6* and *COR78*, respectively (Figure 4.13F-H). The data suggest that reducing or increasing expression of a single stress-responsive gene can alter biomass accumulation. Combinatorial effects of repressing many stress-responsive genes in hybrids could contribute to biomass heterosis.

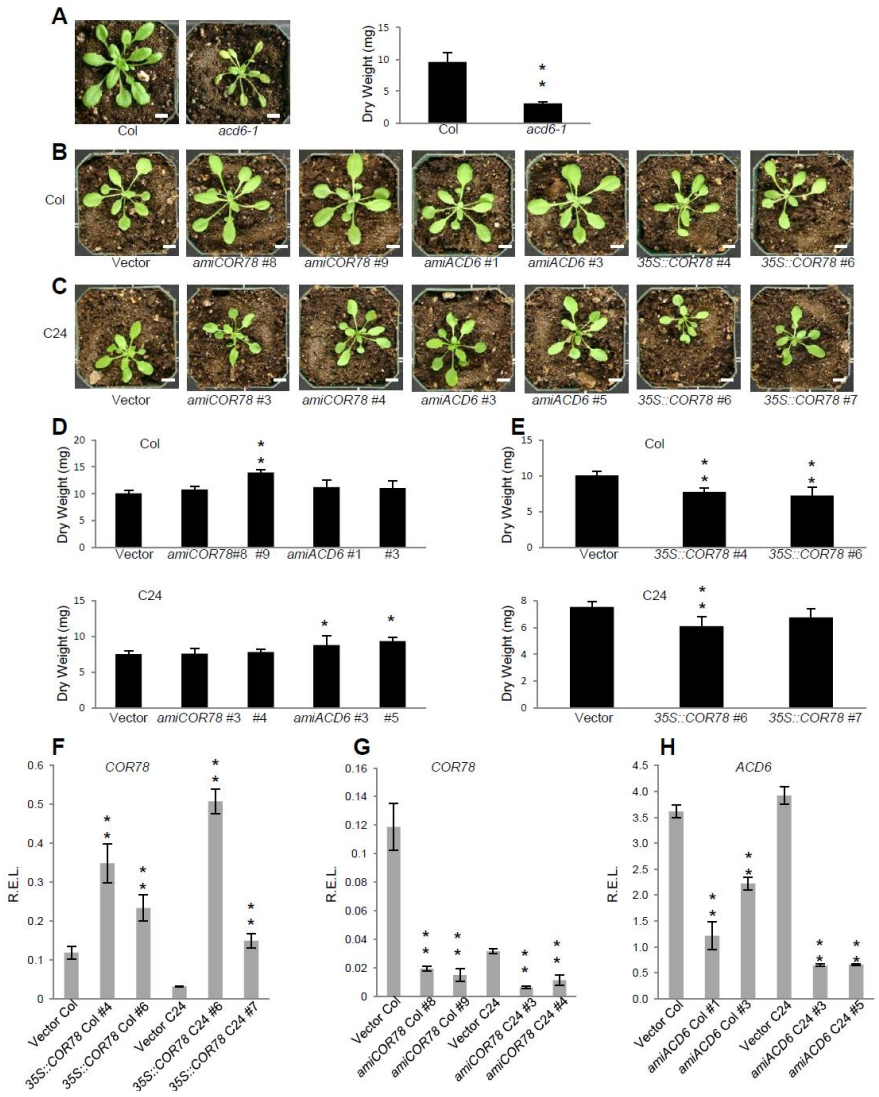


Figure 4.13 Effects of knockdown or overexpression of abiotic and biotic genes on biomass

(**A**) Seedlings of Col and *acd6-1* mutant (left) and their biomass (right). (**B**) Seedlings of transgenic Col that expressed vector alone, *amiRCOR78*, *amiRACD6*, or *35S:COR78*. (**C**) Seedlings of transgenic C24 that expressed vector alone, *amiRCOR78*, *amiRACD6*, or *35S:COR78*. Scale bars = 1 cm. (**D**) Dry weight (mg) of transgenic *amiRCOR78* or *amiRACD6* lines in Col (upper) and in C24 (lower). (**E**) Dry weight (mg) of transgenic *35S:COR78* lines in Col (upper) and in C24 (lower). (**F-G**) R.E.L of *COR78* expression in overexpression lines (**F**) and in *amiRNA* lines (**G**). (**H**) R.E.L. of *ACD6* in amiRNA lines. Values were averaged from three biological replicates (± s.d.). Single and double asterisks indicate statistical significance levels at $p<0.05$ and $p<0.01$ using two-tailed student's t-test, respectively, compared to the transgenic control (vector).

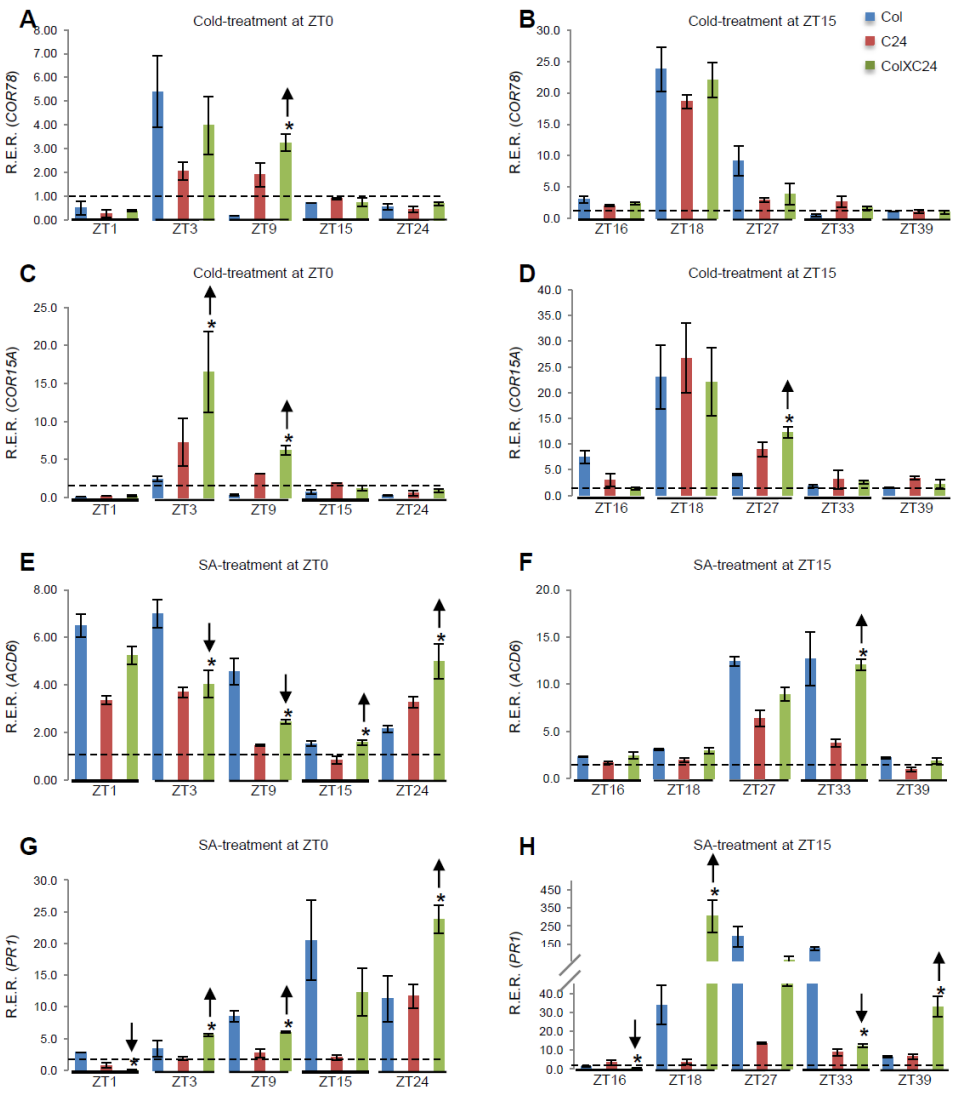


Figure 4.14 The induction of stress-responsive gene expression in F1(ColXC24) hybrids and their parents (Col and C24)

Relative expression ratios (R.E.R.) of *COR78* (**A, B**) and *COR15A* (**C, D**) between treated and untreated samples across 5 time points after the cold-treatment at ZT0 (**A, C**) or ZT15 (**B, D**). R.E.R. of *ACD6* (**E, F**) and *PR1* (**G, H**) between treated and untreated samples across 5 time points after the SA-treatment at ZT0 (**E, G**) or ZT15 (**F, H**). Expression values were averaged from three biological replicates (± s.d.). Asterisks indicate statistical significance levels at $p<0.05$ using two-tailed student's t-test, compared to the MPV. Dashed lines indicate no expression change between treated and untreated samples. After seedlings were cold- or SA-treated for 1 hour at ZT0 or ZT15, rosette leaves were harvested at designated time points for gene expression studies (see the Methods for details).

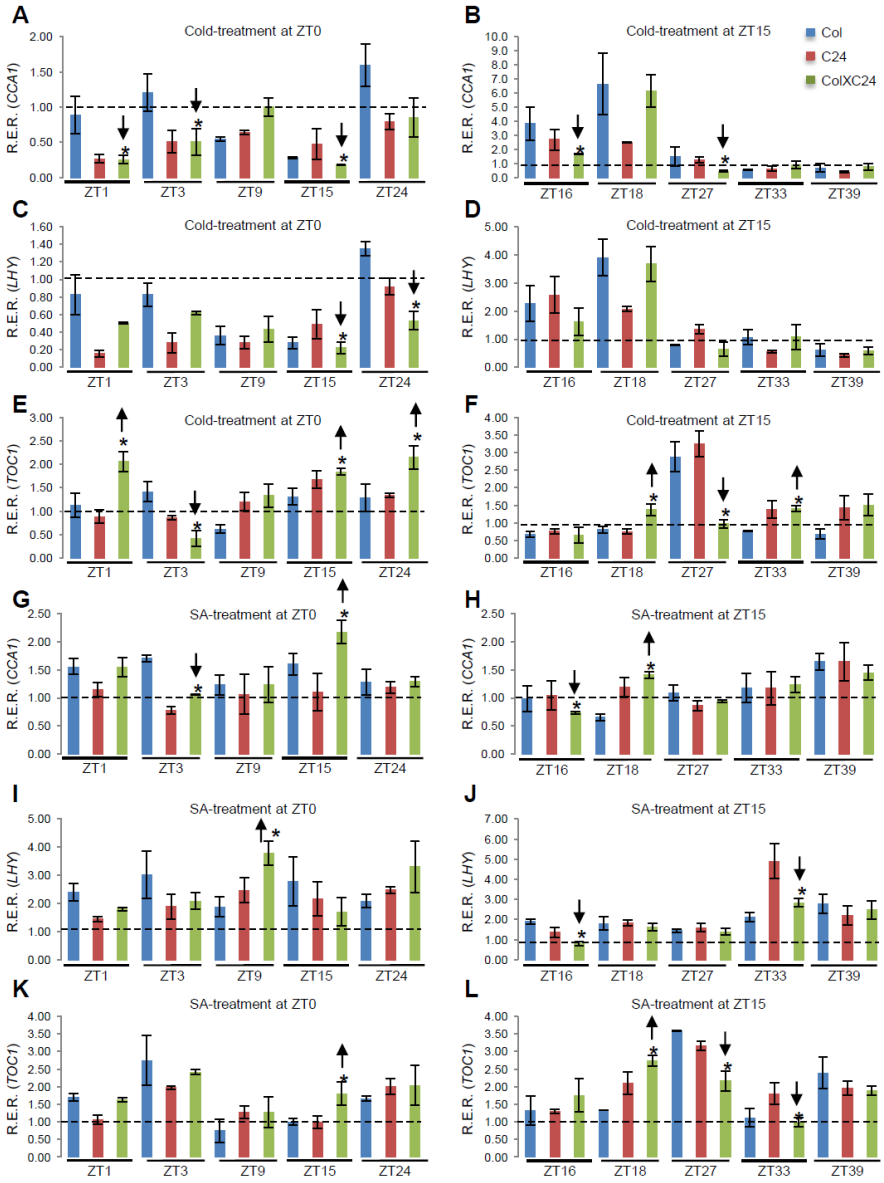


Figure 4.15 Altered expression of circadian clock genes after stress-treatment

R.E.R. of *CCA1* (**A, B, G, H**), *LHY* (**C, D, I, J**), and *TOC1* (**E, F, K, L**) between treated and untreated samples across 5 time points after the indicated stress-treatment at ZT0 (**A, C, E, G, I, K**) or ZT15 (**B, D, F, H, J, L**). Expression values were averaged from three biological replicates (± s.d.). Asterisks indicate statistical significance levels at $p<0.05$ using two-tailed student's t-test, compared to the mid-parent value (MPV). Dashed lines indicate no expression change between treated and untreated samples. After seedlings were cold- or SA-treated for 1 hour at ZT0 or ZT15, rosette leaves were harvested at designated time points for gene expression studies (see the Methods for details).

A role for timing of stress-responsive gene induction in biomass heterosis

It is paradoxical that hybrids show repression of stress-responsive genes in the absence of stress, and yet they are more tolerant under stress conditions [99, 100]. To resolve this paradox, induction of stress-responsive genes was examined in the hybrids and their parents in a 24-hour period under stress conditions. Although a previous study [100] measured the induction of *COR* genes in hybrids and parents after cold-treatment, they neglected to measure multiple points and they did not report the time of day at which they did measure expression. To gain a more nuanced view of gene expression induction in hybrids, cold or SA treatment was applied at either ZT0 or ZT15, and gene expression was examined at subsequent time points (see Methods). After cold shock, both *COR78* and *COR15A* transcripts were induced in ColXC24 hybrids and parents during the day (Figure 4.14A-D). Cold-treatment caused higher fold-increase of gene expression at ZT15 than at ZT0 (Figure 4.14A vs. B and C vs. D). Interestingly, the hybrids reached higher than MPV at some specific time points (e.g., *COR15A* at ZT3 and ZT9 and *COR78* at ZT9). After SA-treatment, both *ACD6* and *PR1* were also induced to higher levels at ZT15 than at ZT0 (Figure 4.14E-H). In the hybrids, induction of *PR1* and *ACD6* was delayed several hours after the treatments, but ultimately reached higher than MPV at several time points.

This temporal regulation of stress-responsive gene expression in the hybrids correlated with altered expression of circadian clock genes. After the cold shock at ZT0, *CCA1* and *LHY* expression decreased before reaching control levels (Figure 4.15A,C), while *TOC1* expression was moderately increased (Figure 4.15E). Opposite changes in the expression of *CCA1* and *LHY* (Figure 4.15B,D) and *TOC1* (Figure 4.15E,F) were also observed after cold shock at ZT15. After the SA treatment, *CCA1*, *LHY*, and *TOC1* transcripts were slightly elevated, with similar levels of expression changes observed at

ZT0 and ZT15 (Figure 4.15G-L). These data suggest a link between altered expression of circadian clock genes and differential induction of stress-responsive genes in the hybrids.

To determine if hybrids perform better under stress conditions, relative growth rates (RGR) were quantified, and final biomass was examined (see Methods). Under cold stress, hybrids maintained a larger rosette size than the parents. After removal from stress conditions the RGR of hybrids accelerated more than that of the parents, and hybrids maintained a higher biomass relative to the parents growing in the presence of stress (Figures 4.16A,B,E and 4.17A,B). After two weeks growing in the cold, hybrids accumulated more biomass relative to the parents compared to hybrids growing in the absence of stress (Figure 4.16E). After two weeks growing in the cold, hybrids also showed higher induction of cold-responsive genes at certain times of day (Figure 4.17C,D).

The growth rate increase between hybrids and parents was more evident under SA treatment than under cold stress (Figures 4.16EC,D and 4.17E,F). Dry weight was statistically significantly more increased in the hybrids relative to the parents after the SA treatment than in the hybrids relative to the parents in the control (Figure 4.16F). After two applications of SA over a two week period (Methods), transcript levels of stress-responsive genes were not as elevated in the hybrids as in the parents (Figure 4.17G,H). Hybrids did not constantly maintain higher than mid-parent values of some biotic stress-responsive genes under prolonged stress. The data collectively suggest that stress-responsive genes, which are normally repressed in the hybrids, can in some cases be rapidly induced to higher than MPV levels at certain time points in response to stresses, preserving energy and metabolism for growth in the absence of stress.

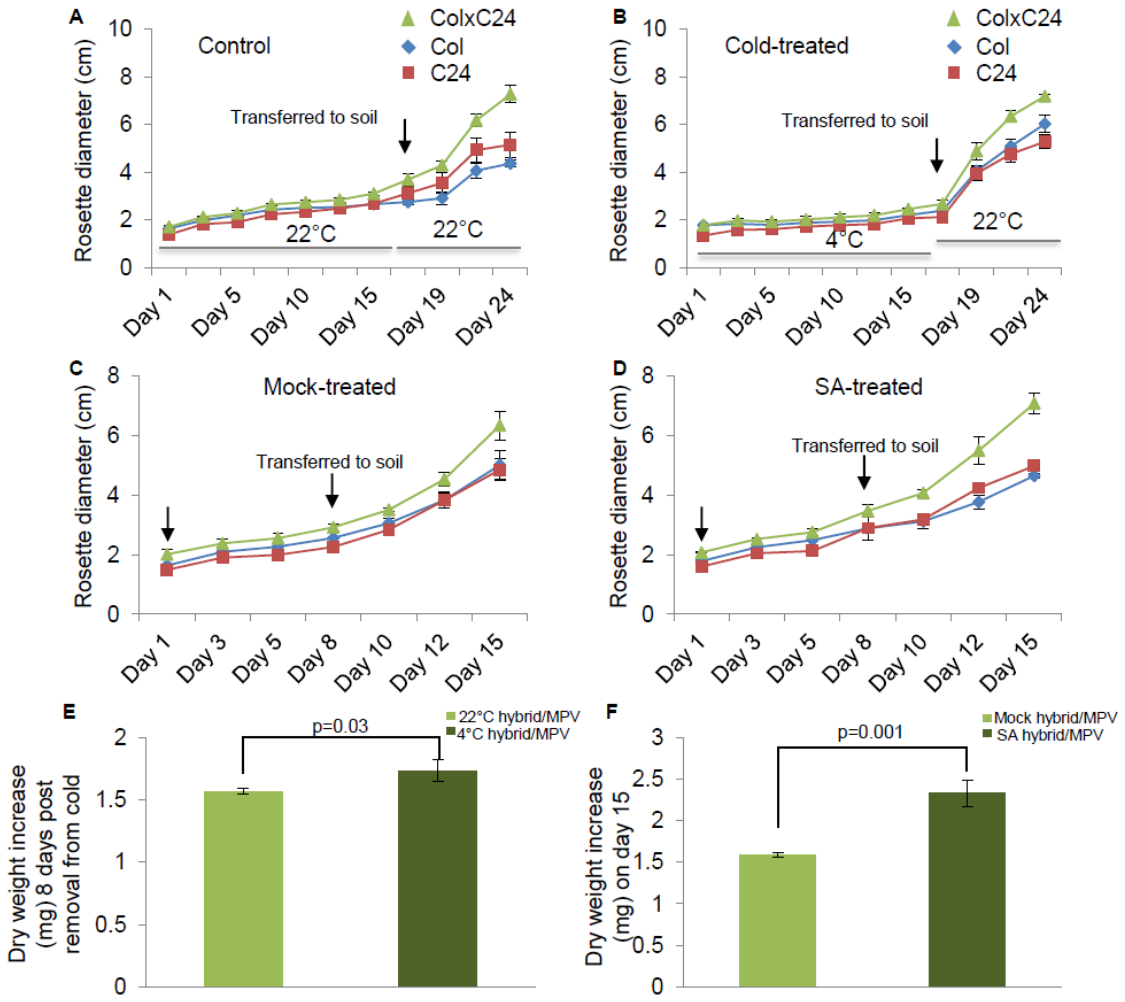


Figure 4.16 Effects of cold stress and salicylic acid (SA) on the growth rate and biomass in hybrids and their parents

(**A, B**) Rosette diameters of seedlings (n = 15 plants averaged per replicate) that were treated at 22ºC (**A**) or 4ºC (**B**) for 17 days and transferred to grow in soil for another 8 days at 22ºC. (**C, D**) Rosette diameters of seedlings (n = 15 plants averaged per replicate) that were mock treated (**C**) or treated with salicylic acid twice (indicated by arrows) (SA) (**D**) and were transferred to soil for another 8 days. (**E**) Dry weight (n = 5 plants averaged per replicate) of the hybrids relative to the MPV in the control and cold-treated conditions 8 days post transfer. (**J**) Dry weight (n = 5 plants averaged per replicate) of the hybrids relative to the MPV in the control and SA-treated conditions after 15 days. Values were averaged from three biological replicates (± s.d.). See the Methods for detailed descriptions. Single and double asterisks indicate statistically significant levels of $p < 0.05$ and $p<0.01$ using two-tailed student's t-test.

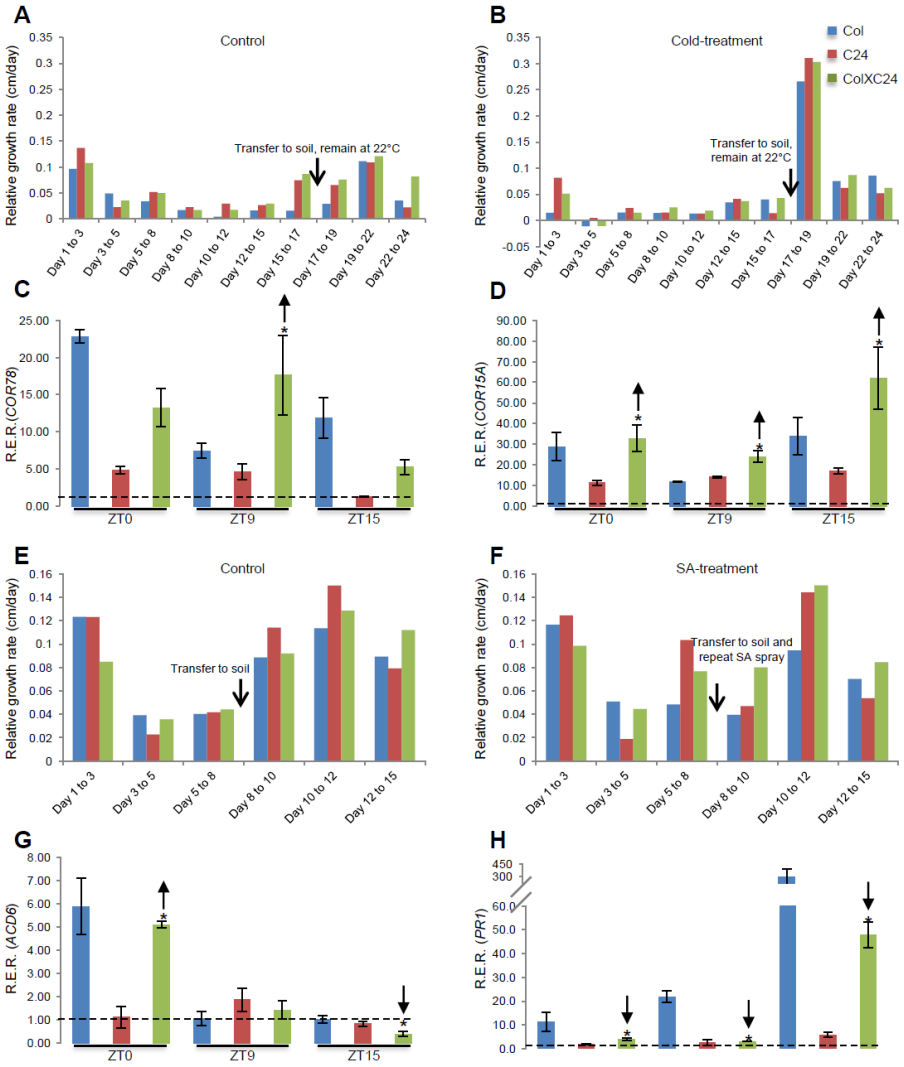


Figure 4.17 Longer-term cold stress and salicylic acid (SA) affect seedling growth rate and gene expression in F1 (ColXC24) hybrids and their parents

Relative growth rates (rosette diameters, cm) of control (untreated) (**A, E**) and treated (**B, F**) seedlings. R.E.R. of *COR78* (**C**), *COR15A* (**D**), *ACD6* (**G**), and *PR1* (h) between treated and control samples across 3 time points. Dashed lines indicate no expression change between treated and untreated samples. Expression values are averaged from three biological replicates (± s.d.). Asterisks indicate statistical significance levels at $p<0.05$ using two-tailed student's t-test, compared to the mid-parent value (MPV). See the Methods for detailed descriptions.

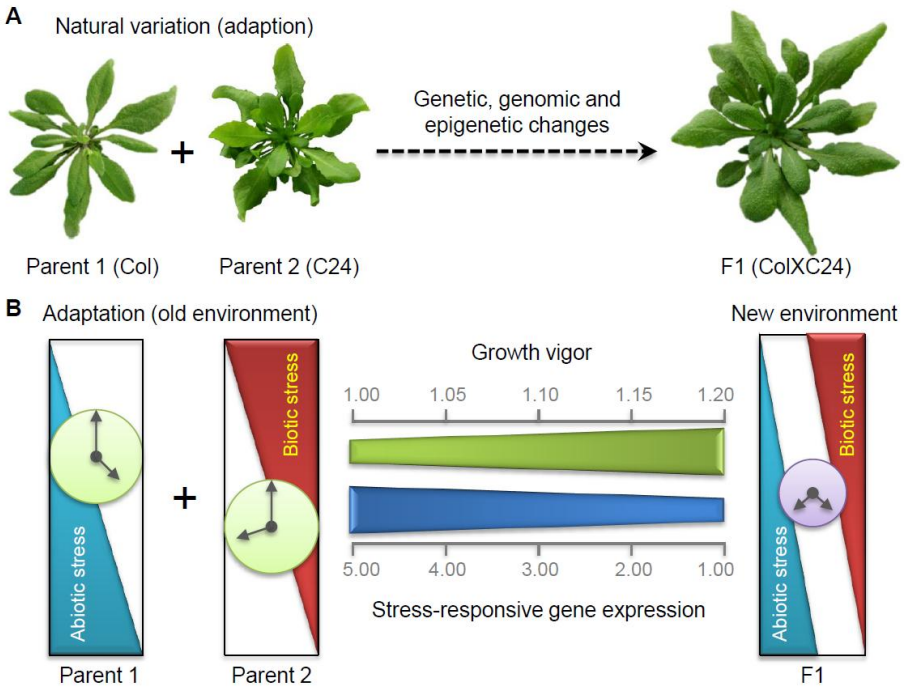


Figure 4.18 A model for how altered stress-responsive gene expression promotes growth vigor in hybrids

(**A**) An example of F1(ColXC24) hybrid showing growth vigor compared to the parents Col and C24 (in the same scale). (**B**) Natural variation at genetic and epigenetic levels is associated with adaptation to ecological niches. Parent 1 and parent 2 could be adapted to different environments that require high levels of gene expression in response to abiotic stress (parent 1) or biotic stress (parent 2), which is partly mediated by circadian rhythms (clock symbols). Parent 1 and parent 2 do not reach the full growth potential probably because of the fitness cost in response to the maintaining high levels of a particular stress response. In F1 hybrids, expression of both biotic and abiotic stress-responsive genes is compromised, probably as consequences of genetic and epigenetic changes and altered circadian rhythms, leading to increased levels of growth vigor. Hybrids have a better stress management plan because the expression of stress-responsive genes is induced at certain times of the day to balance the tradeoff between an urgently needed stress response and long-term growth vigor.

A model for altering stress-responsive gene expression to promote growth vigor in hybrids

The available data support a model for natural variation and circadian regulation of stress-responsive gene expression in heterosis (Figure 4.18). In a given environment, ecotypes (parents) must balance the tradeoff between defense and growth (Figure 4.18A). This adaptive requirement for diurnal and seasonal changes in response to stresses could lead to differences in circadian rhythms. As a result, expression of stress-responsive genes is partly regulated by the circadian clock. Moreover, this "long-term" adaptation to corresponding environments could result in genetic and epigenetic variation between the parents, which is reprogrammed in the hybrids to alter expression of circadian regulators [31, 63] and stress-responsive genes. Indeed, overall levels of CG, CHG and CHH (H = A, T, or C) methylation in many genes, including biotic and abiotic stress-responsive genes (Figure 4.19A,B), were higher in the hybrids than in the parents [31] (Figure 4.19C-K). Increased methylation was associated with small interfering RNAs that accumulated in a region that is adjacent to *PR1* (Figure 4.19A) or downstream in *ACD6* (Figure 4.19B). Consistent with the methylation increase, these genes were repressed in the hybrids (C24XL*er*) at a level similar to the low parent (Figure 4.19A,B). These results are reminiscent of dynamic DNA methylation changes in transposons and repeats that can regulate expression of neighboring genes in response to biotic stress [178]. In addition to both *CCA1* and *LHY* being upregulated in *ddc* mutants (methyltransferase null mutants), both *ACD6* and *PR1*, in addition to many other stress genes, are also upregulated, indicating that increased methylation levels in hybrids contribute to repression of stress and circadian-regulated genes (see [138] Supplemental Table 3).

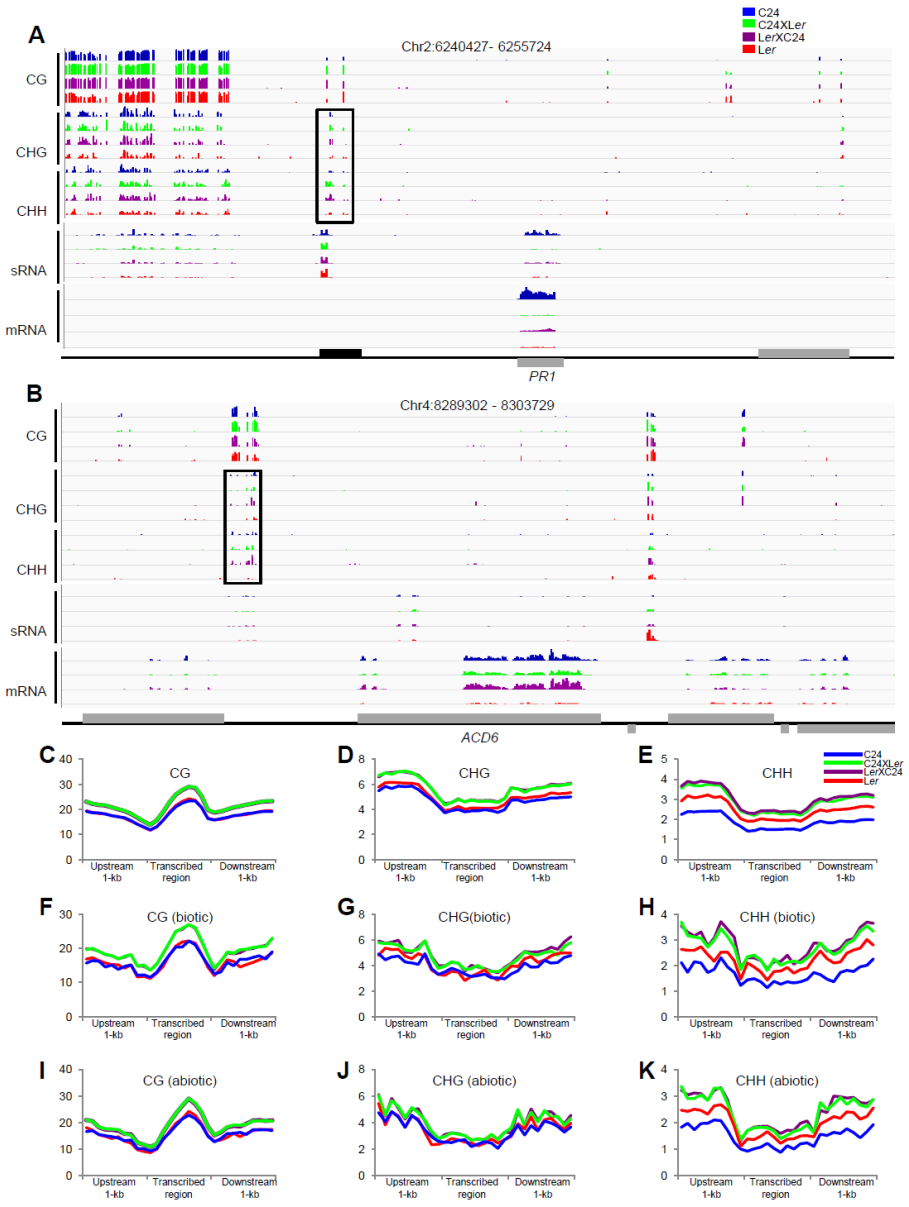


Figure 4.19 Sequence read densities and distributions of DNA methylation, small RNA, and mRNA in the vicinity of *PR1* and *ACD6* genomic regions

Gray boxes are genes, and black boxes are transposons. Changes in CHG and CHH methylation between F1 reciprocal hybrids and their parents (L*er* and C24) are boxed. Genomic coordinates are shown above each diagram. Percentage (%, Y-axis) of CG (**C**), CHG (**D**), and CHH (**E**) methylation levels in all genes. Percentage (%, Y-axis) of CG (**F**), CHG (**G**), and CHH (**H**) methylation levels in all biotic stress responsive genes. Percentage (%, Y-axis) of CG (**I**), CHG (**J**), and CHH (**K**) methylation levels in all abiotic stress responsive genes. X-axis: positions relative to the transcribed region. Biotic and abiotic gene lists were obtained from genome annotation version 10 at the Arabidopsis Information Resource (TAIR). Data were extracted from the published paper (Shen et al. Plant Cell, 2012).

Altered expression of stress-responsive genes could consequently increase or decrease growth vigor. On one hand, this model could explain low levels of heterosis or hybrid necrosis as reported [179], where the expression of stress-responsive genes such as *PR1*, *PR2*, and *PR5* is increased, rather than decreased, in the hybrids. On the other hand, higher-vigor hybrids have a better stress management plan: repressing stress-responsive genes under normal conditions and inducing stress responses at certain times of the day during times of stress, thus balancing the tradeoff between a rapid requirement for stress responses and long-term maintenance of growth vigor (Figure 4.18B). This type of stress management plan is inherited from divergent ecotypes of *A. thaliana*, with more competitive species having lower levels of constitutive expression of stress-responsive genes but higher levels of inducible resistance [174]. Hybrids have simply exploited this adaptive response. In mice, young hybrids also survive better than inbred parents in cold conditions [180]. In plants, hybrids could recover faster than the parents after encountering stress, suggesting that hybrids have a higher potential to promote growth by regulating stress responses in both normal and stress conditions. This concept of linking natural variation and circadian regulation of stress-responsive genes with heterosis could guide the selection of the parents to improve hybrid production in plants and animals.

References

1. Darwin, C., *The effects of cross and self fertilisation in the vegetable kingdom*. 2d ed. 1878, London,: J. Murray. viii, 487, 1 p. incl. tables.
2. Shull, G.H., *The composition of a field of maize.* Am. Breeders Assoc. Rep., 1908. **4**: p. 296-301.
3. East, E.M., *Reports of the Connecticut Agricultural Experiment Station for Years 1907-1908.* Connecticut Agricultural Experiment Station, 1908: p. 419-428
4. Chen, Z.J., *Molecular mechanisms of polyploidy and hybrid vigor.* Trends Plant Sci, 2010. **15**(2): p. 57-71.
5. Mallet, J., *Hybridization as an invasion of the genome.* Trends Ecol Evol, 2005. **20**(5): p. 229-37.
6. Schnable, P.S. and N.M. Springer, *Progress toward understanding heterosis in crop plants.* Annu Rev Plant Biol, 2013. **64**: p. 71-88.
7. Lippman, Z.B. and D. Zamir, *Heterosis: revisiting the magic.* Trends Genet, 2007. **23**(2): p. 60-6.
8. Barth, S., et al., *Heterosis for biomass yield and related traits in five hybrids of Arabidopsis thaliana L. Heynh.* Heredity, 2003. **91**(1): p. 36-42.
9. Meyer, R.C., et al., *Heterosis of biomass production in Arabidopsis. Establishment during early development.* Plant Physiol, 2004. **134**(4): p. 1813-23.
10. Nordborg, M., et al., *The pattern of polymorphism in Arabidopsis thaliana.* PLoS Biol, 2005. **3**(7): p. e196.
11. Hannah, M.A., et al., *Natural genetic variation of freezing tolerance in Arabidopsis.* Plant Physiol, 2006. **142**(1): p. 98-112.
12. Atwell, S., et al., *Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines.* Nature, 2010. **465**(7298): p. 627-31.
13. Groszmann, M., et al., *Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor.* Proc Natl Acad Sci U S A, 2011.
14. Chen, Z.J., *Genomic and epigenetic insights into the molecular bases of heterosis.* Nat Rev Genet, 2013. **14**(7): p. 471-482.
15. East, E.M., *Heterosis.* Genetics, 1936. **21**(4): p. 375-97.
16. Duvick, D.N., *Biotechnology in the 1930s: the development of hybrid maize.* Nat Rev Genet, 2001. **2**(1): p. 69-74.
17. Jones, D.F., *Dominance of Linked Factors as a Means of Accounting for Heterosis.* Genetics, 1917. **2**(5): p. 466-79.
18. Birchler, J.A., D.L. Auger, and N.C. Riddle, *In search of the molecular basis of heterosis.* Plant Cell, 2003. **15**(10): p. 2236-9.
19. Duvick, D.N., *Heterosis: Feeding people and protecting natural resources.* Genetics and Exploitation of Heterosis in Crops, 1999: p. 19-29.

20. Li, Z.K., et al., *Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield.* Genetics, 2001. **158**(4): p. 1737-53.
21. Xiao, J., et al., *Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers.* Genetics, 1995. **140**(2): p. 745-54.
22. Birchler, J.A., et al., *Heterosis.* Plant Cell, 2010. **22**(7): p. 2105-12.
23. Song, R. and J. Messing, *Gene expression of a gene family in maize based on noncollinear haplotypes.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 9055-60.
24. Guo, M., et al., *Allelic variation of gene expression in maize hybrids.* Plant Cell, 2004. **16**(7): p. 1707-16.
25. Auger, D.L., et al., *Nonadditive gene expression in diploid and triploid hybrids of maize.* Genetics, 2005. **169**(1): p. 389-97.
26. Swanson-Wagner, R.A., et al., *All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6805-10.
27. Springer, N.M. and R.M. Stupar, *Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize.* Plant Cell, 2007. **19**(8): p. 2391-402.
28. He, G., et al., *Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids.* Genome Biol, 2013. **14**(6): p. R57.
29. He, G., et al., *Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids.* Plant Cell, 2010. **22**(1): p. 17-33.
30. Fujimoto, R., et al., *Heterosis of Arabidopsis hybrids between C24 and Col is associated with increased photosynthesis capacity.* Proc Natl Acad Sci U S A, 2012.
31. Shen, H., et al., *Genome-Wide Analysis of DNA Methylation and Gene Expression Changes in Two Arabidopsis Ecotypes and Their Reciprocal Hybrids.* Plant Cell, 2012.
32. Wang, J., et al., *Genomewide nonadditive gene regulation in Arabidopsis allotetraploids.* Genetics, 2006. **172**(1): p. 507-17.
33. Wang, J., et al., *Stochastic and epigenetic changes of gene expression in Arabidopsis polyploids.* Genetics, 2004. **167**(4): p. 1961-73.
34. Comai, L., et al., *Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids.* Plant Cell, 2000. **12**(9): p. 1551-68.
35. Gaeta, R.T., et al., *Genomic changes in resynthesized Brassica napus and their effect on gene expression and phenotype.* Plant Cell, 2007. **19**(11): p. 3403-17.
36. Marmagne, A., et al., *Analysis of gene expression in resynthesized Brassica napus allotetraploids: transcriptional changes do not explain differential protein regulation.* New Phytol, 2010. **186**(1): p. 216-27.
37. Adams, K.L., et al., *Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4649-54.

38. Flagel, L., et al., *Duplicate gene expression in allopolyploid Gossypium reveals two temporally distinct phases of expression evolution.* BMC Biol, 2008. **6**: p. 16.
39. Hegarty, M.J., et al., *Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication.* Curr Biol, 2006. **16**(16): p. 1652-9.
40. Chelaifa, H., A. Monnier, and M. Ainouche, *Transcriptomic changes following recent natural hybridization and allopolyploidy in the salt marsh species Spartina x townsendii and Spartina anglica (Poaceae).* New Phytol, 2010. **186**(1): p. 161-74.
41. Tate, J.A., et al., *Evolution and expression of homeologous loci in Tragopogon miscellus (Asteraceae), a recent and reciprocally formed allopolyploid.* Genetics, 2006. **173**(3): p. 1599-611.
42. Buggs, R.J., et al., *Gene loss and silencing in Tragopogon miscellus (Asteraceae): comparison of natural and synthetic allotetraploids.* Heredity (Edinb), 2009. **103**(1): p. 73-81.
43. Pumphrey, M., et al., *Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat.* Genetics, 2009. **181**(3): p. 1147-57.
44. Chague, V., et al., *Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids.* New Phytol, 2010. **187**(4): p. 1181-94.
45. Dahal, D., B.P. Mooney, and K.J. Newton, *Specific changes in total and mitochondrial proteomes are associated with higher levels of heterosis in maize hybrids.* Plant J, 2012. **72**(1): p. 70-83.
46. Hoecker, N., et al., *Analysis of nonadditive protein accumulation in young primary roots of a maize (Zea mays L.) F(1)-hybrid compared to its parental inbred lines.* Proteomics, 2008. **8**(18): p. 3882-94.
47. Marcon, C., et al., *Nonadditive protein accumulation patterns in Maize (Zea mays L.) hybrids during embryo development.* J Proteome Res, 2010. **9**(12): p. 6511-22.
48. Wang, W., et al., *Proteomic profiling of rice embryos from a hybrid rice cultivar and its parental lines.* Proteomics, 2008. **8**(22): p. 4808-21.
49. Meyer, R.C., et al., *Heterosis manifestation during early Arabidopsis seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids.* Plant J, 2012. **71**(4): p. 669-83.
50. Riedelsheimer, C., et al., *Genomic and metabolic prediction of complex heterotic traits in hybrid maize.* Nat Genet, 2012. **44**(2): p. 217-20.
51. Feher, K., et al., *Deducing hybrid performance from parental metabolic profiles of young primary roots of maize by using a multivariate diallel approach.* PLoS One, 2014. **9**(1): p. e85435.
52. Greaves, I.K., et al., *Trans Chromosomal Methylation in Arabidopsis hybrids.* Proc Natl Acad Sci U S A, 2012. **109**(9): p. 3570-5.
53. Greaves, I., et al., *Trans-chromosomal methylation.* Epigenetics, 2012. **7**(8): p. 800-5.

54. Moghaddam, A.M., et al., *Additive inheritance of histone modifications in Arabidopsis thaliana intra-specific hybrids.* Plant J, 2011. **67**(4): p. 691-700.
55. Eichten, S.R., et al., *Heritable epigenetic variation among maize inbreds.* PLoS Genet, 2011. **7**(11): p. e1002372.
56. Greaves, I.K., et al., *Inheritance of Trans Chromosomal Methylation patterns from Arabidopsis F1 hybrids.* Proc Natl Acad Sci U S A, 2014.
57. Chodavarapu, R.K., et al., *Transcriptome and methylome interactions in rice hybrids.* Proc Natl Acad Sci U S A, 2012. **109**(30): p. 12040-5.
58. Li, Y., et al., *The inheritance pattern of 24 nt siRNA clusters in arabidopsis hybrids is influenced by proximity to transposable elements.* PLoS One, 2012. **7**(10): p. e47043.
59. Ha, M., et al., *Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids.* Proc Natl Acad Sci U S A, 2009. **106**(42): p. 17835-40.
60. Kenan-Eichler, M., et al., *Wheat hybridization and polyploidization results in deregulation of small RNAs.* Genetics, 2011. **188**(2): p. 263-72.
61. Barber, W.T., et al., *Repeat associated small RNAs vary among parents and following hybridization in maize.* Proc Natl Acad Sci U S A, 2012. **109**(26): p. 10444-9.
62. He, G., H. He, and X.W. Deng, *Epigenetic variations in plant hybrids and their potential roles in heterosis.* J Genet Genomics, 2013. **40**(5): p. 205-10.
63. Ni, Z., et al., *Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids.* Nature, 2009. **457**(7227): p. 327-31.
64. Kim, E.D. and Z.J. Chen, *Unstable transcripts in Arabidopsis allotetraploids are associated with nonadditive gene expression in response to abiotic and biotic stresses.* PLoS One, 2011. **6**(8): p. e24251.
65. Ha, M., E.D. Kim, and Z.J. Chen, *Duplicate genes increase expression diversity in closely related species and allopolyploids.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2295-300.
66. Ng, D.W., et al., *Proteomic divergence in Arabidopsis autopolyploids and allopolyploids and their progenitors.* Heredity, 2011.
67. Covington, M.F., et al., *Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development.* Genome Biol, 2008. **9**(8): p. R130.
68. McClung, C.R., *Plant circadian rhythms.* Plant Cell, 2006. **18**(4): p. 792-803.
69. Dodd, A.N., et al., *Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage.* Science, 2005. **309**(5734): p. 630-3.
70. Wijnen, H. and M.W. Young, *Interplay of circadian clocks and metabolic rhythms.* Annu Rev Genet, 2006. **40**: p. 409-48.
71. Bass, J. and J.S. Takahashi, *Circadian integration of metabolism and energetics.* Science, 2010. **330**(6009): p. 1349-54.
72. Prasai, M.J., J.T. George, and E.M. Scott, *Molecular clocks, type 2 diabetes and cardiovascular disease.* Diab Vasc Dis Res, 2008. **5**(2): p. 89-95.

73. Turek, F.W., et al., *Obesity and metabolic syndrome in circadian Clock mutant mice.* Science, 2005. **308**(5724): p. 1043-5.
74. Graf, A., et al., *Circadian control of carbohydrate availability for growth in Arabidopsis plants at night.* Proc Natl Acad Sci U S A, 2010. **107**(20): p. 9458-63.
75. Miller, M., C.Q. Zhang, and Z.J. Chen, *Ploidy and Hybridity Effects on Growth Vigor and Gene Expression in Arabidopsis thaliana Hybrids and Their Parents.* G3-Genes Genomes Genetics, 2012. **2**(4): p. 505-513.
76. Song, G.S., et al., *Comparative transcriptional profiling and preliminary study on heterosis mechanism of super-hybrid rice.* Mol Plant, 2010. **3**(6): p. 1012-25.
77. Dong, M.A., E.M. Farre, and M.F. Thomashow, *Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis.* Proc Natl Acad Sci U S A, 2011. **108**(17): p. 7241-6.
78. Legnaioli, T., J. Cuevas, and P. Mas, *TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought.* EMBO J, 2009. **28**(23): p. 3745-57.
79. Wang, W., et al., *Timing of plant immune responses by a central circadian regulator.* Nature, 2011. **470**(7332): p. 110-4.
80. Goodspeed, D., et al., *Arabidopsis synchronizes jasmonate-mediated defense with insect circadian behavior.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(12): p. 4674-4677.
81. Zhang, C., et al., *Crosstalk between the Circadian Clock and Innate Immunity in Arabidopsis.* PLoS Pathog, 2013. **9**(6): p. e1003370.
82. Bhardwaj, V., et al., *Defence responses of Arabidopsis thaliana to infection by Pseudomonas syringae are regulated by the circadian clock.* PLoS One, 2011. **6**(10): p. e26968.
83. Covington, M.F. and S.L. Harmer, *The circadian clock regulates auxin signaling and responses in Arabidopsis.* PLoS Biol, 2007. **5**(8): p. e222.
84. Yakir, E., et al., *Regulation of output from the plant circadian clock.* FEBS J, 2007. **274**(2): p. 335-45.
85. Rawat, R., et al., *REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways.* Proc Natl Acad Sci U S A, 2009. **106**(39): p. 16883-8.
86. Harmer, S.L., et al., *Orchestrated transcription of key pathways in Arabidopsis by the circadian clock.* Science, 2000. **290**(5499): p. 2110-3.
87. Kreps, J.A., et al., *Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress.* Plant Physiol, 2002. **130**(4): p. 2129-41.
88. Fowler, S.G., D. Cook, and M.F. Thomashow, *Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock.* Plant Physiol, 2005. **137**(3): p. 961-8.
89. Bieniawska, Z., et al., *Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome.* Plant Physiol, 2008. **147**(1): p. 263-79.

90. Kidokoro, S., et al., *The phytochrome-interacting factor PIF7 negatively regulates DREB1 expression under circadian control in Arabidopsis.* Plant Physiol, 2009. **151**(4): p. 2046-57.
91. Goodspeed, D., et al., *Postharvest circadian entrainment enhances crop pest resistance and phytochemical cycling.* Curr Biol, 2013. **23**(13): p. 1235-41.
92. Bryant, P.A., J. Trinder, and N. Curtis, *Sick and tired: Does sleep have a vital role in the immune system?* Nat Rev Immunol, 2004. **4**(6): p. 457-67.
93. Tian, D., et al., *Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana.* Nature, 2003. **423**(6935): p. 74-7.
94. Kasuga, M., et al., *Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor.* Nat Biotechnol, 1999. **17**(3): p. 287-91.
95. Heil, M. and I.T. Baldwin, *Fitness costs of induced resistance: emerging experimental support for a slippery concept.* Trends Plant Sci, 2002. **7**(2): p. 61-7.
96. Burdon, J.J. and P.H. Thrall, *The fitness costs to plants of resistance to pathogens.* Genome Biol, 2003. **4**(9): p. 227.
97. Dai, X., et al., *Overexpression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis.* Plant Physiol, 2007. **143**(4): p. 1739-51.
98. Todesco, M., et al., *Natural allelic variation underlying a major fitness trade-off in Arabidopsis thaliana.* Nature, 2010. **465**(7298): p. 632-6.
99. Korn, M., et al., *Heterosis in the freezing tolerance, and sugar and flavonoid contents of crosses between Arabidopsis thaliana accessions of widely varying freezing tolerance.* Plant Cell Environ, 2008. **31**(6): p. 813-27.
100. Rohde, P., D.K. Hincha, and A.G. Heyer, *Heterosis in the freezing tolerance of crosses between two Arabidopsis thaliana accessions (Columbia-0 and C24) that show differences in non-acclimated and acclimated freezing tolerance.* Plant J, 2004. **38**(5): p. 790-9.
101. Saibo, N.J., T. Lourenco, and M.M. Oliveira, *Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses.* Ann Bot, 2009. **103**(4): p. 609-23.
102. Bilgin, D.D., et al., *Biotic stress globally downregulates photosynthesis genes.* Plant Cell Environ, 2010. **33**(10): p. 1597-613.
103. Riddle, N.C., et al., *Gene expression analysis at the intersection of ploidy and hybridity in maize.* Theor Appl Genet, 2010. **120**(2): p. 341-53.
104. Riddle, N.C., A. Kato, and J.A. Birchler, *Genetic variation for the response to ploidy change in Zea mays L.* Theor Appl Genet, 2006. **114**(1): p. 101-11.
105. Lysak, M., P. Fransz, and I. Schubert, *Cytogenetic analyses of Arabidopsis.* Methods Mol Biol, 2006. **323**: p. 173-86.
106. Galbraith, D.W., et al., *Rapid flow cytometric analysis of the cell cycle in intact plant tissues.* Science, 1983. **220**(4601): p. 1049-51.

107. Dart, S., P. Kron, and B.K. Mable, *Characterizing polyploidy in Arabidopsis lyrata using chromosome counts and flow cytometry.* Canadian Journal of Botany-Revue Canadienne De Botanique, 2004. **82**(2): p. 185-197.
108. Bell, C.J. and J.R. Ecker, *Assignment of 30 microsatellite loci to the linkage map of Arabidopsis.* Genomics, 1994. **19**(1): p. 137-44.
109. Jeffrey Chen, Z., et al., *The development of an Arabidopsis model system for genome-wide analysis of polyploidy effects.* Biol J Linn Soc Lond, 2004. **82**(4): p. 689-700.
110. Galbraith, D.W., K.R. Harkins, and S. Knapp, *Systemic Endopolyploidy in Arabidopsis thaliana.* Plant Physiol, 1991. **96**(3): p. 985-9.
111. McFadden, D.E., et al., *Parental origin of triploidy in human fetuses: evidence for genomic imprinting.* Hum Genet, 1993. **92**(5): p. 465-9.
112. Nadeau, J.A., *Stomatal development: new signals and fate determinants.* Current Opinion in Plant Biology, 2009. **12**(1): p. 29-35.
113. McClung, C.R., *A modern circadian clock in the common angiosperm ancestor of monocots and eudicots.* BMC Biol, 2010. **8**: p. 55.
114. Pruneda-Paz, J.L. and S.A. Kay, *An expanding universe of circadian networks in higher plants.* Trends Plant Sci, 2010. **15**(5): p. 259-65.
115. Reinbothe, S., et al., *PORA and PORB, Two Light-Dependent Protochlorophyllide-Reducing Enzymes of Angiosperm Chlorophyll Biosynthesis.* Plant Cell, 1996. **8**(5): p. 763-769.
116. Henry, I.M., et al., *Phenotypic consequences of aneuploidy in Arabidopsis thaliana.* Genetics, 2010. **186**(4): p. 1231-45.
117. Ravi, M. and S.W. Chan, *Haploid plants produced by centromere-mediated genome elimination.* Nature, 2010. **464**(7288): p. 615-8.
118. Masterson, J., *Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms.* Science, 1994. **264**(5157): p. 421-4.
119. Beaulieu, J.M., et al., *Genome size is a strong predictor of cell size and stomatal density in angiosperms.* New Phytol, 2008. **179**(4): p. 975-86.
120. Bretagnolle, F., J.D. Thompson, and R. Lumaret, *The Influence of Seed Size Variation on Seed-Germination and Seedling Vigor in Diploid and Tetraploid Dactylis-Glomerata L.* Annals of Botany, 1995. **76**(6): p. 607-615.
121. Alonso-Blanco, C., et al., *Natural allelic variation at seed size loci in relation to other life history traits of Arabidopsis thaliana.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4710-7.
122. Ashby, E., *Studies in the inheritance of physiological characters. III. hybrid vigour in the tomato. Part 1. manifestation of hybrid vigour from germination to the onset of flowering.* Ann Bot (Lond), 1937. **1**: p. 11–41.
123. Haig, D. and M. Westoby, *Parent-Specific Gene-Expression and the Triploid Endosperm.* American Naturalist, 1989. **134**(1): p. 147-155.
124. Scott, R.J., et al., *Parent-of-origin effects on seed development in Arabidopsis thaliana.* Development, 1998. **125**(17): p. 3329-41.

125. Tiwari, S., et al., *Transcriptional profiles underlying parent-of-origin effects in seeds of Arabidopsis thaliana.* BMC Plant Biol, 2010. **10**: p. 72.
126. Moore, T. and D. Haig, *Genomic imprinting in mammalian development: a parental tug-of-war.* Trends Genet, 1991. **7**(2): p. 45-9.
127. Tilghman, S.M., *The sins of the fathers and mothers: genomic imprinting in mammalian development.* Cell, 1999. **96**(2): p. 185-93.
128. Dilkes, B.P. and L. Comai, *A differential dosage hypothesis for parental effects in seed development.* Plant Cell, 2004. **16**(12): p. 3174-80.
129. Finn, T.E., et al., *Transgene expression and transgene-induced silencing in diploid and autotetraploid Arabidopsis.* Genetics, 2011. **187**(2): p. 409-23.
130. Onodera, Y., et al., *Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation.* Cell, 2005. **120**(5): p. 613-22.
131. Haag, J.R. and C.S. Pikaard, *Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing.* Nat Rev Mol Cell Biol, 2011. **12**(8): p. 483-92.
132. Wassenegger, M., et al., *RNA-directed de novo methylation of genomic sequences in plants.* Cell, 1994. **76**(3): p. 567-76.
133. Law, J.A. and S.E. Jacobsen, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals.* Nat Rev Genet, 2010. **11**(3): p. 204-20.
134. Herr, A.J., et al., *RNA polymerase IV directs silencing of endogenous DNA.* Science, 2005. **308**(5718): p. 118-20.
135. Zilberman, D., et al., *Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats.* Curr Biol, 2004. **14**(13): p. 1214-20.
136. Gao, Z., et al., *An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation.* Nature, 2010. **465**(7294): p. 106-9.
137. Chan, S.W., I.R. Henderson, and S.E. Jacobsen, *Gardening the genome: DNA methylation in Arabidopsis thaliana.* Nat Rev Genet, 2005. **6**(5): p. 351-60.
138. Kurihara, Y., et al., *Identification of the candidate genes regulated by RNA-directed DNA methylation in Arabidopsis.* Biochem Biophys Res Commun, 2008. **376**(3): p. 553-7.
139. Kankel, M.W., et al., *Arabidopsis MET1 cytosine methyltransferase mutants.* Genetics, 2003. **163**(3): p. 1109-22.
140. Jeddeloh, J.A., J. Bender, and E.J. Richards, *The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis.* Genes Dev, 1998. **12**(11): p. 1714-25.
141. Salome, P.A. and C.R. McClung, *PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock.* Plant Cell, 2005. **17**(3): p. 791-803.
142. McGinnis, K., et al., *Transgene-induced RNA interference as a tool for plant functional genomics.* Methods Enzymol, 2005. **392**: p. 1-24.

143. Clough, S.J. and A.F. Bent, *Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana.* Plant Journal, 1998. **16**(6): p. 735-743.
144. Gruntman, E., et al., *Kismeth: analyzer of plant methylation states through bisulfite sequencing.* BMC Bioinformatics, 2008. **9**: p. 371.
145. Huh, J.H., et al., *Cellular programming of plant gene imprinting.* Cell, 2008. **132**(5): p. 735-44.
146. Ferguson-Smith, A.C., *Genomic imprinting: the emergence of an epigenetic paradigm.* Nat Rev Genet, 2011. **12**(8): p. 565-75.
147. Raissig, M.T., C. Baroux, and U. Grossniklaus, *Regulation and flexibility of genomic imprinting during seed development.* Plant Cell, 2011. **23**(1): p. 16-26.
148. Zilberman, D., et al., *Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription.* Nat Genet, 2007. **39**(1): p. 61-9.
149. Pruneda-Paz, J.L., et al., *A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock.* Science, 2009. **323**(5920): p. 1481-5.
150. James, A.B., et al., *The circadian clock in Arabidopsis roots is a simplified slave version of the clock in shoots.* Science, 2008. **322**(5909): p. 1832-5.
151. Penfield, S. and A. Hall, *A role for multiple circadian clock genes in the response to signals that break seed dormancy in Arabidopsis.* Plant Cell, 2009. **21**(6): p. 1722-32.
152. Nakahata, Y., et al., *Circadian control of the NAD+ salvage pathway by CLOCK-SIRT1.* Science, 2009. **324**(5927): p. 654-7.
153. Mosher, R.A., et al., *Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis.* Nature, 2009. **460**(7252): p. 283-6.
154. Lu, J., et al., *Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds.* Proc Natl Acad Sci U S A, 2012. **109**(14): p. 5529-34.
155. Cao, X., et al., *Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation.* Curr Biol, 2003. **13**(24): p. 2212-7.
156. Zemach, A., et al., *The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin.* Cell, 2013. **153**(1): p. 193-205.
157. Harmer, S.L., *The circadian system in higher plants.* Annu Rev Plant Biol, 2009. **60**: p. 357-77.
158. Nagel, D.H. and S.A. Kay, *Complexity in the wiring and regulation of plant circadian networks.* Curr Biol, 2012. **22**(16): p. R648-57.
159. Haig, D., *Coadaptation and conflict, misconception and muddle, in the evolution of genomic imprinting.* Heredity (Edinb), 2013.
160. van Wees, S., *Phenotypic Analysis of Arabidopsis Mutants: Trypan Blue Stain for Fungi, Oomycetes, and Dead Plant Cells.* Cold Spring Harbor Protocols, 2008. **2008**(8): p. pdb.prot4982.

161. Homer, N., B. Merriman, and S.F. Nelson, *BFAST: an alignment tool for large scale genome resequencing.* PLoS One, 2009. **4**(11): p. e7767.
162. Shi, X.L., et al., *Cis- and trans-regulatory divergence between progenitor species determines gene-expression novelty in Arabidopsis allopolyploids.* Nature Communications, 2012. **3**.
163. Wang, L., et al., *DEGseq: an R package for identifying differentially expressed genes from RNA-seq data.* Bioinformatics, 2010. **26**(1): p. 136-8.
164. Michael, T.P. and C.R. McClung, *Phase-specific circadian clock regulatory elements in Arabidopsis.* Plant Physiol, 2002. **130**(2): p. 627-38.
165. Dodd, A.N., et al., *Time of day modulates low-temperature Ca signals in Arabidopsis.* Plant J, 2006. **48**(6): p. 962-73.
166. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nature Methods, 2012. **9**(7): p. 671-675.
167. van Hulten, M., et al., *Costs and benefits of priming for defense in Arabidopsis.* Proceedings of the National Academy of Sciences, 2006. **103**(14): p. 5602-5607.
168. Michael, T.P., et al., *A morning-specific phytohormone gene expression program underlying rhythmic plant growth.* PLoS Biol, 2008. **6**(9): p. e225.
169. Streitner, C., et al., *Global transcript profiling of transgenic plants constitutively overexpressing the RNA-binding protein AtGRP7.* BMC Plant Biol, 2010. **10**: p. 221.
170. Korn, M., et al., *Predicting Arabidopsis freezing tolerance and heterosis in freezing tolerance from metabolite composition.* Mol Plant, 2010. **3**(1): p. 224-35.
171. Bechtold, U., et al., *Constitutive salicylic acid defences do not compromise seed yield, drought tolerance and water productivity in the Arabidopsis accession C24.* Plant Cell and Environment, 2010. **33**(11): p. 1959-1973.
172. Lalic, J., et al., *Adaptation of tobacco etch potyvirus to a susceptible ecotype of Arabidopsis thaliana capacitates it for systemic infection of resistant ecotypes.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2010. **365**(1548): p. 1997-2007.
173. Jones, J.D. and J.L. Dangl, *The plant immune system.* Nature, 2006. **444**(7117): p. 323-9.
174. Kempel, A., et al., *Tradeoffs associated with constitutive and induced plant resistance against herbivory.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(14): p. 5685-5689.
175. Orgil, U., et al., *Intraspecific genetic variations, fitness cost and benefit of RPW8, a disease resistance locus in Arabidopsis thaliana.* Genetics, 2007. **176**(4): p. 2317-2333.
176. Rate, D.N., et al., *The gain-of-function Arabidopsis acd6 mutant reveals novel regulation and function of the salicylic acid signaling pathway in controlling cell death, defenses, and cell growth.* Plant Cell, 1999. **11**(9): p. 1695-1708.
177. Gilmour, S.J., et al., *Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation.* Plant Physiology, 2000. **124**(4): p. 1854-1865.

178. Dowen, R.H., et al., *Widespread dynamic DNA methylation in response to biotic stress.* Proceedings of the National Academy of Sciences, 2012. **109**(32): p. E2183–E2191.
179. Bomblies, K., et al., *Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants.* PLoS Biol, 2007. **5**(9): p. e236.
180. Barnett, S.A., *Heterozygosis and the Survival of Young Mice in Two Temperatures.* Q J Exp Physiol Cogn Med Sci, 1964. **49**: p. 290-6.

Vita

Marisa Elena Miller was born in Austin, Texas. After graduating from Gonzalo Garza Independence High School in 2003, she attended Austin Community College from 2004-2006. Subsequently, she entered the University of Texas at Austin and received her Bachelor of Science in Cell and Molecular Biology in 2008. In September of 2008, she entered the Cell and Molecular Biology graduate program at the University of Texas at Austin where she joined the lab of Dr. Z. Jeffrey Chen in May 2009.

Permanent e-mail address: marisa.e.miller@gmail.com

This dissertation was typed by Marisa Elena Miller.